2015

Robust Healthcare Financing Systems

Tengjiao Xiao
Lehigh University

Robust Healthcare Financing Systems

by

Tengjiao Xiao

Presented to the Graduate and Research Committee

of Lehigh University

in Candidacy for the Degree of

Doctor of Philosophy

in

Industrial Engineering

Lehigh University

May 2015

Approved and recommended for acceptance as a dissertation in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

_____
Date

_____
Dissertation Advisor

Committee Members:

_____
Dr. Aurélie C. Thiele, Committee Chair

_____
Dr. Boris Defourny

_____
Dr. Ruken Duzgun

_____
Stuart Paxton

_____
Dr. Luis Zuluaga

Acknowledgments

I would like to express the deepest gratitude to my advisor, Dr. Aurélie C. Thiele, for her persistent encouragement and support. This thesis would not have been possible without her generous help and guidance in my research and career. I sincerely appreciate her patience and kindness and her way of "always thinking of students first". I am lucky to have her as my advisor.

I would also like to thank my committee members, Dr. Boris Defourny, Dr. Ruken Duzgun, Dr. Luis Zuluaga for serving on my thesis committee and giving valuable comments on this thesis. A special thanks to Professor Stuart Paxton, who shared with me his insights in the healthcare industry and gave me the opportunity to apply my knowledge and skills to the real business world. I am also grateful to acknowledge the support from the department, Rossin Doctoral Fellowship, and travel grant from Graduate Student Senate.

This thesis is dedicated to my loving parents, who are always being there for me whenever I need and give me a sweet home. I also want to express my great appreciation to my uncle, aunt and grandma for giving me all they could have given when my parents are not around. Last but not the least, Xi, thank you for your unconditional love, support, and company. I am grateful for who you are, and I am grateful for what you do.

Contents

List of Figures

List of Tables

Abstract

This dissertation provides robust, quantitative models in healthcare finance to aid decision-makers with rigorous, analytical tools that capture high complexity and high uncertainty of problem. The first chapter investigates the impact of parameter uncertainty on risk scoring, and presents an approach to obtain robust risk scores to address uncertainty in risk adjustment, which is used to quantify payment transfers across health plans under the Affordable Care Act. We provide a tractable methodology to incorporate uncertainty in the risk factor weights via linear programming to improve risk adjustment among payers and discuss the impact of uncertainty on the risk scores. In the second chapter, we provide an analytical methodology to help individuals narrow down plan choices in the Health Insurance Exchanges (HIX) by identifying plans that are dominated by competitors' offerings in terms of premium, metal level, maximum out of pocket payment and plan type. We further quantify the amount by which the premium of a dominated plan should be reduced to make it competitive in our framework. This part of our work provides important quantitative tools to guide the discussions between payers and policy-makers regarding HIX. Our approach also provides payers with a novel way to analyze their own plans in the HIX landscape. The third chapter identifies the key factors that drive enrollment rates of the two major types of Medicare Part D prescription drug plans: MAPD and PDP, to assist policy makers in better promoting their plans to Medi-

care beneficiaries. The fourth chapter investigates trends in physician services usage and Medicare reimbursement rate from CMS public files. We analyze the HCPCS (Healthcare Common Procedure Coding System) codes and investigate the validity of the concern that doctors tend to upcode on purpose for more reimbursement. We also utilize time series analysis to predict Medicare spending in ten years. In the last chapter, we survey and propose robust optimization models in healthcare systems engineering, particularly in the applications of healthcare costs prediction, disease management, IMRT fluence map optimization, and operating room planning, among others.

Chapter 1

Contributions

The health care system in the United States is complex and health care costs are growing at unsustainable rates. In 2013, U.S. health care spending reached $2.9 trillion, an average of $9,255 per person. It accounted for 17.4 percent of GDP of that year. Although the United States has the most expensive health care system in the world, the Commonwealth Fund [30] ranked the United States last in the quality of health care among seven developed countries - Australia, Canada, Germany, the Netherlands, New Zealand, and the United Kingdom - on measures of health system performance in five areas: quality, efficiency, access to care, equity and the ability to lead long, healthy, productive lives. In order to improve the health care system, President Obama signed comprehensive health reform, the Patient Protection and Affordable Care Act, into law on March 23, 2010. The ultimate goals of healthcare reform are to expand coverage, control health care costs, and improve health care delivery system. However, according to preliminary data released on April 30 by the Bureau of Economic Analysis, total health care spending during the first quarter of 2014 increased by 9.9% to $43.3 billion and contributed 1.1 percentage points to growth in the total gross domestic product. The purpose of this dissertation is

to provide robust, quantitative models in healthcare finance to aid decision-makers with rigorous, analytical tools that capture high complexity and high uncertainty of problem. The contributions of each chapter are as follows.

We investigate healthcare financing systems from five different perspectives: risk adjustment under uncertainty to help payers be properly compensated for the health status of their enrollees, the efficiency of health insurance plans offered on public exchanges to help customers get the most value out of their money, Medicare enrollees' choice regarding Medicare Advantage and stand-alone prescription drug plans to ensure long-term affordability of prescription drugs for an aging population, the trends in physician services usage for the Medicare population and Medicare reimbursement rates, and the use of robust optimization techniques to protect decision-makers under uncertainty in a wide range of health systems engineering settings.

Regarding risk adjustment under uncertainty, our contributions are: (1) We discuss the impact of parameter uncertainty on risk scoring and risk adjustment; (2) We present a highly tractable approach to create robust risk scores to incorporate ambiguity and uncertainty in risk adjustment model. The highly tractable approach is based on solving linear programming problems; (3) We provide empirical results on the impact of robust risk scoring on actual money transfer for insurers.

Regarding the efficiency of health insurance plans on HIX, our contributions are: (1) We investigate the main drivers of plan premiums in the health exchanges, with examples drawn from the Pennsylvania and Massachusetts health insurance marketplaces; (2) We propose a simple algorithm based on linear regression to identify inefficient (dominated) plans in the exchanges based on plans' attributes, and to assign letter grades to payers; (3) We identify the excess premium by which certain plans are less competitive than others based on an analysis of plans' features such

as premiums and deductibles.

On Medicare beneficiaries' choice between Medicare Advantage and stand-alone prescription drug plans, our contributions are: (1) We provide an overview of the major two types of Medicare Part D prescription drug plans: Medicare Advantage prescription plans (MAPDs), and stand-alone prescription plans (PDPs); (2) We identify the key factors that drive enrollment rates of MAPDs and PDPs in all counties in the United States using beta regression, and discuss the differences in factors driving MAPD and PDP enrollment; (3) We make recommendations to assist policy makers in better promoting their plans to Medicare beneficiaries.

On the trends in physician services usage for the Medicare population and Medicare reimbursement rates, our contributions are: (1) We investigate trends in physician services usage and Medicare reimbursement rate from CMS public files; (2) We analyze the HCPCS (Healthcare Common Procedure Coding System) codes and investigate the validity of the concern that doctors tend to upcode on purpose to get more reimbursement; (3) We utilize time series analysis and linear regression to predict Pennsylvania's Medicare spending in ten years.

On the applications of robust optimization techniques to protect decision-makers under uncertainty in a wide range of health systems engineering settings, our contributions are: (1) We provide an overview of the healthcare systems, including payment systems and delivery systems, in the United States; (2) We survey and propose robust optimization models in healthcare systems engineering, particularly in the applications of healthcare costs prediction, disease management, IMRT fluence map optimization, and operating room planning, among others.

Chapter 2

Robust Risk Adjustment in Health Insurance

2.1 Introduction

Risk adjustment is defined in the Specifications Manual for National Hospital Quality Measures as "a statistical process used to identify and adjust for variation in patient outcomes that stem from differences in patient characteristics (or risk factors) across health care organizations." [77] The goal of risk adjustment is to reflect that "patients may experience different outcomes regardless of the quality of care provided by the health care organization" due to patient-specific characteristics, such as age or clinical diagnoses [77]. Without appropriate risk adjustment, comparing patient outcomes across organizations can be misleading. For instance, a best-in-class health provider may attract particularly ill patients, who may face dire prognoses and thus may have worse outcomes than patients who are only moderately ill and go to a less-skilled provider. By accounting for existing risk factors, risk adjustment facilitates

a more fair and accurate inter-organizational comparison. The broad concepts and applications of risk adjustment are presented in Ellis [35].

Risk adjustment is further defined by the American Academy of Actuaries as "an actuarial tool used to calibrate payments to health plans or other stakeholders based on the relative health of the at-risk populations." [1] In that context, it extends beyond risk measurement into risk mitigation, and helps ensure that health plans are appropriately compensated for the risks they enroll. Risk adjustment is a permanent "zero-sum game" in that the total amounts paid by health plans into the risk adjustment pool are received by other health plans each year. The goal of this program is to stabilize a competitive marketplace in which health plans compete on plan features and services rather than on avoidance of high risk individuals. This is beneficial to consumers, particularly those with high-cost health conditions, as it is more likely to give them continued choice of health plans. Specifically, risk adjustment – when done well – can help remove the incentive for health plans to try not to enroll sicker people (aka adverse selection), since they will be compensated for those patients' worse health status. Adverse selection can lead to three classes of inefficiencies: prices to participants do not reflect marginal costs, hence on a benefit-cost basis individuals select the wrong health plans; desirable risk spreading is lost; and health plans manipulate their offerings to deter the sick and attract the healthy [29]. Since insurers set premiums based on the riskiness of the people they enroll, adverse selection would also lead to higher premiums and government spending [17]. Glazer et. al. [50] develop a statistical methodology to improve upon adverse selection outcome in design of risk adjustment formula in health insurance markets, where enrollees sort between plans with fixed benefit offerings as a function of the plans' premiums. McWilliams et. al. [61] show using a regression model that the implementation of the Hierarchical Condition Categories (HCC) model was associated with

reduced favorable selection in the Medicare Advantage program. However, they also point out that inadequate risk adjustment would probably cause greater instability in exchange markets than in Medicare Advantage, and lead to competition among exchange plans to attract and retain healthy enrollees, as well as the withdrawal of undercompensated plans. In contrast, Brown et. al. [19] demonstrate that the effect of risk adjustment on government's cost of providing health insurance is imperfect since risk adjustment can potentially increase the scope for selecting individuals with costs below their capitation payment due to the increase in the variance of medical costs with the risk score.

The Health and Human Services (HHS) official federal risk adjustment models are available in the HHS Notice of Benefit and Payment Parameters for 2014, which was first released as a proposed rule at the end of 2012 [75]. The HHS risk adjustment system uses fifteen weighted least squares regression models: platinum, gold, silver, bronze, and catastrophic for adult, child, and infant, respectively, to compute risk scores. The weight is the fraction of the year enrolled. Each HHS risk adjustment model predicts annual plan liability for an enrollee based on the person's age, gender, and diagnoses. The risk score of each enrollee is equal to the sum of all the risk weights associated with that patient, with the average risk score over the whole population being scaled to 1.

The enrollment-weighted average risk score of all enrollees in a particular health plan within a geographic rating area (the weights being again the fractions of the year enrolled) are then used as input to the payment transfer formula to determine an issuer's payment or charge for a particular plan, which is a baseline payment times the plan's enrollment-weighted average risk score [5]. The HHS risk adjustment model is a concurrent model, where diagnoses from a given period are used to predict cost in the same period. In contrast, a prospective model uses data from

a prior period to predict costs in the current period or in the future. By design, both acute and chronic illnesses are emphasized in the concurrent model. In the prospective model, systematic factors, such as aging and chronic illnesses, outweigh acute and one-time conditions [2]. Acute and one-time events are averaged at the age/gender group level in the prospective model (Yi et. al. [91]). The concurrent model is used by HHS because it is more robust to changes in enrollment than the prospective model ([91], [87]). This is particularly useful under the Affordable Care Act since newly enrolled individuals may not have prior claims data. In addition, prescription drugs are not included as a predictor in each HHS risk adjustment model. To evaluate model performance, R^2 and predictive ratios are examined, where the R^2 statistic calculates the percentage of individual variation explained by a model, and the predictive ratio is the ratio of the weighted-mean predicted plan liability for the model sample population to the weighted-mean actual plan liability for the model sample population [91].

Winkelman [88] uses the Mean Absolute Prediction Error (MAPE) as an alternative to measure predictive accuracy, where MAPE is calculated by dividing the sum of absolute errors by the sample size. Glazer and McGuire [49] argue that, in order to address adverse selection and asymmetric information in managed care, risk adjustment should be viewed as a way to set prices for different individuals. They argue for instance that the payment weight on a patient's age "may be chosen for its incentive properties and need not – indeed should not – be the same as the coefficient on age from a regression explaining average costs." Weiner et. al. [84] quantify the impact of biased selection on health plans in the exchange and evaluates mitigation attempts included in the Affordable Care Act of 2010.

Proper risk adjustment is thus very important for payers' long-term financial viability and for the competitiveness of the health insurance market. Risk adjustment

has been used in the Medicare Advantage (MA) program, the Part D prescription drug program, many state Medicaid programs, the Commonwealth Care program in Massachusetts, and some employer-based plans [1]. Risk adjustment for commercial insurance arrived in 2014 for the individual and small-group marketplaces. The main difference between CMS-HCC model for Medicare and HHS-HCC model for commercial insurance is that insurers get payments from CMS directly under Medicare, while payments are between insurers under commercial risk adjustment. The HHS methodology was applied to all non-grandfathered plans in all states except Massachusetts both inside and outside of the marketplaces. The risk weights can be obtained by linear regression, probit regression, or logistic regression, depending on the situation considered; however, estimates of regression coefficients are subject to error. Because risk adjustment in this context involves real money transfers between payers, it is important to develop quantitative methods to incorporate ambiguity and uncertainty in the risk weights. The main contribution of this chapter is to present a tractable methodology to create robust risk scores, which determine the amount of money to transfer. While we propose another way to investigate and incorporate impact of uncertainty, out methodology serves as a supplement to current implementation of HHS-HCC model and leads to fair payments when uncertainty exists.

2.2 Calculation of Risk Adjustment in Healthcare

Risk adjustment is typically done by computing risk scores for each enrollee and assigning to each health plan an amount of money equal to a baseline payment weighted by the aggregate risk score for the population it covers [5]. This is because risk scores are computed such that their average is 1. While a risk score attempts to

quantify how "costly" an enrollee will be to a plan in the future, a precise measure requires to specify what information will be used in making that determination: concurrent weights emphasize acute conditions, while prospective weights focus on chronic ones [2]. An example of risk score calculation from the Health Affairs Issue Brief is as follows:

Table 2.1: Example of risk score of a patient

Risk Factor	Risk Weight
Male age 32	0.22
Diabetes w/ significant comorbidities	1.32
Asthma	0.96
Low-cost dermatology	0.30
Total	2.80

In this example, this particular individual has 4 risk factors including 1 demographic factor and 3 diagnosis factors. Different risk weights are assigned to the factors. A higher risk weight means potential higher chance of incurring healthcare cost. By summing up all risk weights, the risk score for this individual is 2.8. The risk score means that this individual is expected to have healthcare costs 2.80 times higher than population average. To get the dollar amount of this individual's healthcare costs, we can simply multiply his total risk score by the population average cost. In addition, the average risk score for a health plan can be calculated as the arithmetic mean of risk scores of all enrollees in the plan.

2.3 Robustness in Risk Adjustment Models

An example to show the need for robustness is based on the Hospital Value-Based Purchasing (VBP) program, established by the Centers for Medicare and Medicaid Services (CMS) [74]. It aims at realigning hospitals' financial incentives by reward-

ing those that provide highest-quality care, with quality of care being quantified as the weighted sum of three sets of measures: process measures (13 measures), patient satisfaction measures (8 measures) and mortality rates for heart attacks, heart failures and pneumonia within 30 days of a patient's leaving the hospital [65]. CMS funds the VBP adjustment scheme by withholding 1% of each hospital's Medicare payments, and re-distributing this pool of money to the hospitals based on the adjustment factors. Hospitals with the lowest adjustment factors receive little to no money back, and thus their 1% of Medicare payments will be lost to them and reassigned to better performing hospitals. Hospitals with the highest adjustment factors receive payments exceeding their initial 1% contribution to the pool. 1% might be ignored by bigger hospitals, but it can have a significant impact on smaller hospitals or hospitals in precarious financial health. Moody's estimates the preliminary median operating margins for non-profit hospitals in FY 2013 to be at 2.2%, a decrease compared to FY 2012 [78]. The significance of the impact of 1% can also be illustrated by the Hospital Readmissions Reduction Program, in which penalties are collected from the hospitals through a percentage reduction in their base Medicare inpatient claims payments, up to a cap. The cap was 1% in fiscal year 2013, when the aggregate amount of penalties was about $280 million against 2217 hospitals.

CMS first issued proxy factors in August 2012 and the actual adjustment factors for these 2,985 hospitals were published by CMS in December 2012, both of which are provided in Table 16 of the FY 2013 Final Rule Tables [76]. The proxy adjustment factors were calculated using historical baseline and performance periods, and would not actually be used to adjust hospital payments. We investigate the variability between proxy and actual scores as follows. We first compute the rank of each hospital, based on the rank of its adjustment factor, with the hospital having the highest (best) adjustment factor receiving rank 1. We then merge the records under

both the proxy and actual systems to compare proxy and actual ranks, as shown in table 2.7. Ranks matter since the program is based on relative performance.

Table 2.2: Proxy vs. Actual VBP Adjustment Factors and Ranks (FY2013)

Proxy Adj. Factor	Actual Adj. Factor	Proxy Rank	Actual Rank
1.0072129779	0.9949792372	21	2887
1.0017818865	0.9920251756	799	2981
1.0035728881	0.9938951967	347	2938
1.0039407695	0.9946071725	282	2915
1.0042747671	0.9958366705	225	2811

The difference in rank is then computed as the proxy rank minus the actual rank, such that a positive difference represents a gain in ranks following the publication of the final (actual) factors. Figure 2.1 shows the differences in ranks from most negative to most positive. Because the total number of hospitals is approximately 3,000, hospitals at the extreme left of the graph represent hospitals that had been expected to perform at the top based on proxy numbers and found themselves at the bottom when the actual numbers were published. Similarly, hospitals at the extreme right represent hospitals that had been deemed at the bottom based on the proxy factors and came out on top with the actual factors. The wide fluctuation between the proxy factors and the actual ones has, to the best of our knowledge, not been discussed in the press or elsewhere. The worst rank loss is a drop of 2,866 spots – from rank 21 to rank 2,887 – by the Meadowview Regional Medical Center in Maysville, KY. The highest gain in rank – from rank 2,659 to rank 144 – is an increase of 2,515 by Loretto Hospital in Chicago, IL. 335 hospitals or 11.81% of the hospitals considered lost 1,000 spots or more and 250 hospitals or 8.81% gained 1,000 or more. The same pattern follows for fiscal year 2014 and fiscal year 2015. Although proxy adjustment factors are not used for payment purpose, they could be used by hospitals in their marketing strategies to negotiate with insurers or attract/maintain customers. The large fluctuations would force them to change strategies after the actual adjustment

factors are revealed. Some may argue there is possibility that it is due to changes in provider quality. However, such large fluctuations over a large percentage of hospitals within a few months' time risk casting doubt on the meaningfulness of the factors and slowing down efforts to move to value-based models, and suggests there is a need to "robustify" factors.

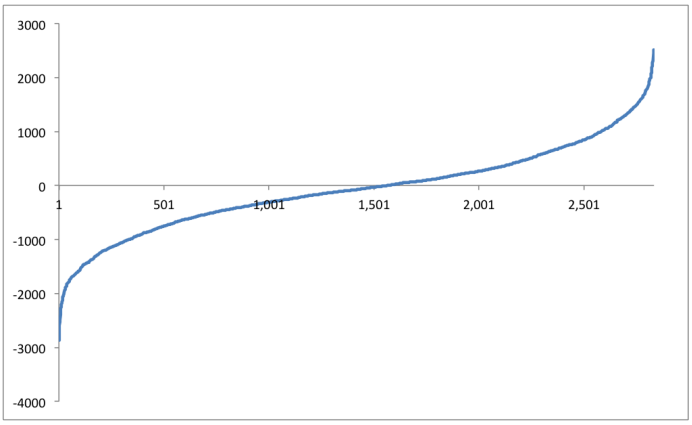


Figure 2.1: Difference in Rank

2.4 Risk Adjustment without Uncertainty

Suppose in the current health insurance market, there are 3 insurers and 1000 patients, each insurer having 1/3 of the patients. We only consider 32 risk factors, including 6 diagnosis factors and 26 demographic factors (age & gender), with risk weights given by the Federal Register Volume 77 Issue 236 (Platinum Plan). The 6 diagnosis factors include chronic diseases such as Asthma, Diabetes, Heart Failure, HIV and Mental Illness, and acute disease Acute Appendicitis. The 26 demographic factors cover individuals aging from 2 to 64. "Made-up" binary parameters for diagnoses and demographics are randomly given to each enrollee.

The traditional risk adjustment process, if the weights of the risk factors are

known exactly, is as follows:

1. Compute the risk score for each enrollee and scale it such that the average population risk score is one,
2. Compute the average risk score for each insurer (weighted by the fraction of the year each enrollee has been on the plan),
3. Determine the transfer payment as the difference between the insurer's cost (sum of patients' risk score times nominal cost) and his revenue (number of patients times capitated payment).

Applying this process to the data we generated, the risk scores of the three insurers would be 0.9861, 1.0190 and 0.9949, respectively. By multiplying the risk scores by the non-adjusted base payment of $500, the transfer payments for each insurer would be -$2,321, $3,163 and -$849, respectively. The sum of all transfer payments are always zero. In this particular example, both insurer 1 and insurer 3 give money, while insurer 2 receives "a lot" of money. However, the impacts on insurer 1 and 3 are quite different since insurer 1 gives "a lot" of money, whereas insurer 3 only gives "a little" money.

In this situation, we assume true values of coefficients are known, which is usually not the case in the real world. Next section we will introduce the situation where uncertainty is being taken into consideration.

2.5 Risk Adjustment with Uncertainty

Suppose risk weights are subject to relative uncertainty with mean 0 and standard deviation 30%. They are truncated to zero if sign changes because risk weights

cannot be negative. We perform 100,000 simulations in MATLAB to quantify the uncertainty. Resulting transfer payments are interpreted as what insurers should have given/received if they had known the true values of the risk weights. Figure 2.2 shows the distribution of transfer payments based on simulation results.

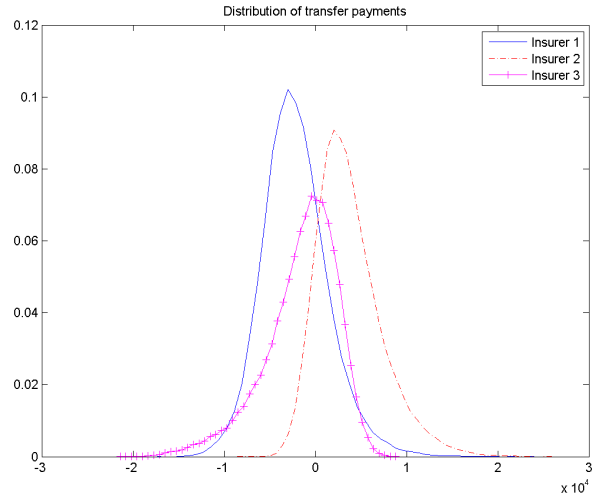


Figure 2.2: Distribution of transfer payments

The distribution of the dataset shows clearly that insurer 1 mostly gives, insurer 2 mostly receives, and insurer 3 is in the middle. Having an idea what the distribution would be like, we look at some basic statistics of the insurers' risk scores from the simulation results.

Table 2.3: Basic Statistics of risk scores

	Insurer 1	Insurer 2	Insurer 3
Expected	0.9870	1.0224	0.9907
St Deviation	0.0217	0.0216	0.0241
Minimum	0.8938	0.9466	0.8695
Maximum	1.1460	1.1577	1.0549

We notice that only insurer 2 has risk score higher than 1, meaning only insurer 2 receives money because it enrolls sicker individuals. The ranking statistics (small-

est/middle/largest risk score) also indicates this conclusion since the probability that insurer 2 will get money is 87.58%.

Table 2.4: Ranking Statistics of risk scores

	Insurer 1	Insurer 2	Insurer 3
Smallest	56.71%	4.91%	38.38%
Middle	31.46%	26.35%	42.19%
Largest	11.83%	68.74%	19.43%
Prob Transfer>0	24.52%	87.58%	39.53%

In addition, we can have the 95% range forecasts for each insurer based on the basic statistics:

Table 2.5: Lower bound and upper bound of risk scores

	Insurer 1	Insurer 2	Insurer 3
Lower Bound	0.9445	0.9801	0.9435
Expected	0.9870	1.0224	0.9907
Upper Bound	1.0295	1.0647	1.0379

Using the expected risk scores, transfer payments in traditional model would be -$2,171, $3,730 and -$1549. However, here the lower bounds and upper bounds of risk scores are not scaled. We scale them to average 1, with loss of the 95% confidence level. Table 2.7 shows the new lower bounds and upper bounds.

Table 2.6: Scaled lower bound and upper bound of risk scores

	Insurer 1	Insurer 2	Insurer 3
“Lower Bound”	0.9879	1.0252	0.9869
“Upper Bound”	0.9861	1.0198	0.9941

Now we obtain scenarios for possible risk scores by creating range forecasts instead of point forecasts for risk scores, with additional constraint that their average is 1:

Table 2.7: Scaled range forecasts of risk scores

	Insurer 1	Insurer 2	Insurer 3
Range	[0.9861,0.9879]	[1.0198,1.0252]	[0.9869,0.9941]

However, this method has some drawbacks. First, these forecasts are obtained by simulation, while we do not know which distribution we should use to simulate the coefficients. Second, scenarios and ranges are obtained by making distributional assumptions (Gaussian distribution) that are hard to check. Third, given ranges, we don't yet know how to transfer money. Therefore, we present a methodology to create robust risk scores to address these issues in the next section.

2.6 Robust Risk Scoring

When the weights for the risk factors are not known precisely but estimates (for instance from a regression) and confidence intervals are available, we face the question of how this uncertainty should be incorporated so that payers receive a "fair" transfer payment. We will seek to minimize the *worst-case regret*. Here the worst-case regret is the greatest difference in absolute value between the estimated and actual risk scores computed over all payers and all possible weights for the risk factors within a predefined uncertainty set. It measures the worst-case difference in absolute value between the money transfer that should have taken place between payers if the true weights had been known and the transfer that actually did, based on the actual weights used to compute the risk scores. These weights are the decision variables of the problem.

We will use the following notation:

K: the number of payers in the market,

S_k: the set of enrollees of insurer $k = 1, \ldots, K$,

J: the set of conditions incorporated in risk scoring,

n_{jk}: the number of enrollees of insurer $k = 1, \ldots, K$ who have condition $j \in J$,

N_k: the number of enrollees of plan k,

c_{ij}: a binary parameter equal to 1 when individual i ($i \in S_k,\ k = 1, \ldots, K$) has condition j,

w_j: the incremental risk weight for condition $j \in J$ (to be added to the risk score of individual i if $c_{ij} = 1$).

Insurer k's risk score before scaling is obtained by taking the average, over all enrollees, of the risk weights of the factors that affect the enrollee.

$$\frac{1}{N_k} \sum_{i \in S_k} \sum_{j \in J} w_j c_{ij} = \frac{1}{N_k} \sum_{j \in J} w_j n_{jk}.$$

For convenience, we assume that all enrollees have been with the payer the whole year. Adapting the formulation to the case where some patients have joined the health plan during the year involves replacing the average over enrollees by a weighted average where the weights are the fractions of year for each patient. This leads to modified definitions for n_{jk} and N_k. Specifically, if τ_{ik} is the fraction of the year individual i has spent with insurer k, n_{jk} becomes $\sum_{i \in S_k} \tau_{ik} c_{ij}$ and N_k becomes $\sum_{i \in S_k} \tau_{ik}$. Once the n_{jk} and N_k have been thus updated, the models presented below apply immediately.

Risk scores are then scaled so that their population average is 1. Insurer k's average risk score after scaling becomes:

$$RS_k = \frac{\sum_{j \in J} w_j n_{jk}}{\sum_{l \in K} \sum_{j \in J} w_j n_{jl}} \cdot \frac{\sum_{l \in K} N_l}{N_k}.$$

We model the uncertain coefficients w_j, $j \in J$, as belonging to a polyhedral set W. The set W can for instance be a box consisting of confidence intervals for each (independent) factor, or possibly include a budget-of-uncertainty constraint in the spirit of Bertsimas and Sim [14] to bound from above the total number of parameters that can take their worst-case value. The problem we aim to solve in the decision variables v (the weights we want to give to each factor within the feasible set W) is then:

$$\min_{v \in W} \max_{k \in K} \max_{w \in W} \left| \frac{\sum_{j \in J} v_j n_{jk}}{\sum_{l \in K} \sum_{j \in J} v_j n_{jl}} \frac{\sum_{l \in K} N_l}{N_k} - \frac{\sum_{j \in J} w_j n_{jk}}{\sum_{l \in K} \sum_{j \in J} w_j n_{jl}} \frac{\sum_{l \in K} N_l}{N_k} \right| \tag{2.1}$$

Let assume w.l.o.g. that the polyhedral set W is represented as $\{w \,|\, l \leq w \leq u,\ Aw = b\}$. Further, let N be the (n_{jk}) matrix, let e be the vector of all ones and let S be the polyhedral set defined as: $\{(x, y) \,|\, ly \leq x \leq uy,\ Ax = b\, y,\ e'N'x = 1\}$. In order to derive a tractable reformulation to Problem (2.1), we will need the following lemma.

Lemma 2.1. *For all $k \in K$, the fractional optimization problems:*

$$(FP\_k): \ \min_{w \in W} \frac{\sum_{j \in J} w_j n_{jk}}{\sum_{l \in K} \sum_{j \in J} w_j n_{jl}},$$

and

$$(FP_+k): \ \max_{w \in W} \frac{\sum_{j \in J} w_j n_{jk}}{\sum_{l \in K} \sum_{j \in J} w_j n_{jl}}.$$

can be solved efficiently by solving the linear programming problems:

$$(LP\_k): \ \min_{(x,y) \in S} \sum_{j \in J} n_{jk} x_j,$$

and

$$(LP_+k): \ \max_{(x,y) \in S} \sum_{j \in J} n_{jk} x_j,$$

respectively.

Proof. The proof is in two steps.

(i) For any $w \in W$, let $x_j = \dfrac{w_j}{\sum_{l\in K}\sum_{j\in J} w_j n_{jl}}$ for all j and $y = \dfrac{1}{\sum_{l\in K}\sum_{j\in J} w_j n_{jl}}$. (Recall that y is always positive because the risk weights and the counts are always positive). Then it is immediate that (x, y) is in the set S defined above and we have: $\dfrac{\sum_{j\in J} w_j n_{jk}}{\sum_{l\in K}\sum_{j\in J} w_j n_{jl}} \dfrac{\sum_{l\in K} N_l}{N_k} = \dfrac{\sum_{l\in K} N_l}{N_k} \cdot \sum_{j\in J} n_{jk} x_j$.

(ii) For any $(x, y) \in S$, we must have $y > 0$ since $y = 0$ would lead to $x = 0$ (due to $ly \le x \le uy$), which would be infeasible (due to $e'N'x = 1$). Let then $w_j = \dfrac{x_j}{y}$ for all j. Then it is immediate that w is in the set W defined above and the two objectives are equal again.

Therefore, Problem (FP_k) is equivalent to (LP_k) and Problem (FP_+k) is equivalent to (LP_+k) for all k. □

Let u_{-k} be the optimal objective of the linear optimization problem (LP_k) and u_{+k} be the optimal objective of the linear optimization problem (LP_+k) for all k. It follows from Lemma 2.1 that we have for all k:

$$u_{-k} = \min_{w\in W} \frac{\sum_{j\in J} w_j n_{jk}}{\sum_{l\in K}\sum_{j\in J} w_j n_{jl}}, \tag{2.2}$$

and

$$u_{+k} = \max_{w\in W} \frac{\sum_{j\in J} w_j n_{jk}}{\sum_{l\in K}\sum_{j\in J} w_j n_{jl}}. \tag{2.3}$$

The key result of this section is the following theorem.

Theorem 2.1 (Robust risk scoring)**.** *Problem (2.1) is equivalent to the following*

linear programming problem:

$$\begin{aligned} \min \quad & Z \\ s.t. \quad & Z \geq \frac{\sum_{l\in K} N_l}{N_k} \left(\sum_{j\in J} n_{jk} x_j - u_{-k} \right) \forall k \\ & Z \geq \frac{\sum_{l\in K} N_l}{N_k} \left(u_{+k} - \sum_{j\in J} n_{jk} x_j \right) \forall k \\ & (x, y) \in S. \end{aligned}$$

Proof. We rewrite Problem (2.1) as:

$$\begin{aligned} \min \quad & Z \\ \text{s.t.} \quad & Z \geq \frac{\sum_{l\in K} N_l}{N_k} \left(\frac{\sum_{j\in J} v_j n_{jk}}{\sum_{l\in K}\sum_{j\in J} v_j n_{jl}} - \frac{\sum_{j\in J} w_j n_{jk}}{\sum_{l\in K}\sum_{j\in J} w_j n_{jl}} \right), \ \forall k, \forall w \in W, \\ & Z \geq \frac{\sum_{l\in K} N_l}{N_k} \left(-\frac{\sum_{j\in J} v_j n_{jk}}{\sum_{l\in K}\sum_{j\in J} v_j n_{jl}} + \frac{\sum_{j\in J} w_j n_{jk}}{\sum_{l\in K}\sum_{j\in J} w_j n_{jl}} \right), \ \forall k, \forall w \in W, \\ & v \in W, \end{aligned}$$

or equivalently:

$$\begin{aligned} \min \quad & Z \\ \text{s.t.} \quad & Z \geq \frac{\sum_{l\in K} N_l}{N_k} \left(\frac{\sum_{j\in J} v_j n_{jk}}{\sum_{l\in K}\sum_{j\in J} v_j n_{jl}} - \min_{w\in W} \frac{\sum_{j\in J} w_j n_{jk}}{\sum_{l\in K}\sum_{j\in J} w_j n_{jl}} \right), \ \forall k, \\ & Z \geq \frac{\sum_{l\in K} N_l}{N_k} \left(-\frac{\sum_{j\in J} v_j n_{jk}}{\sum_{l\in K}\sum_{j\in J} v_j n_{jl}} + \max_{w\in W} \frac{\sum_{j\in J} w_j n_{jk}}{\sum_{l\in K}\sum_{j\in J} w_j n_{jl}} \right), \ \forall k, \\ & v \in W. \end{aligned}$$

We inject Eqs. (2.2) and (2.3) and use the transformation $x_j = \dfrac{v_j}{\sum_{l\in K}\sum_{j\in J} v_j n_{jl}}$ and $y = \dfrac{1}{\sum_{l\in K}\sum_{j\in J} v_j n_{jl}}$. The rest of the proof is identical to the proof of Lemma 2.1, replacing w by v, and is omitted here. □

2.7 Numerical Experiments and Discussions

To test our approach, we generate a sample with 1,000,000 patients and 10 payers. The base payment is $2,000. The risk factors and nominal weights are taken from the Federal Register [75]. The standard error and confidence interval of each risk weight should be available by CMS as they run the regression models. For illustrative purposes, the confidence interval of each risk weight is symmetric, centered at the nominal weight, and with a relative deviation from the mean selected randomly and up to 30% (i.e., the upper bound is at most 1.3 times the nominal weight.) The uncertainty set is a hypercube or "box" consisting of the range forecasts for each weight. Table 2.8 shows the nominal and robust weights as well as the lower and upper bounds of the weights used in the model. Table 2.9 compares nominal and robust risk scores for each insurer.

Table 2.8: Nominal Weights vs. Robust Weights

Risk Factor	Nominal weight	Deviation (%)	Lower bound	Upper bound	Robust weight	Change
Male,21-24	0.258	24.44	0.1949	0.3211	0.1949	-
Male,25-29	0.278	27.17	0.2025	0.3535	0.3533	-/- -
Male,30-34	0.338	3.81	0.3251	0.3509	0.3509	++
Male,35-39	0.413	27.40	0.2998	0.5262	0.2998	- -
Male,40-44	0.487	18.97	0.3946	0.5794	0.3946	- -
Male,45-49	0.581	2.93	0.5640	0.5980	0.5641	-/- -
Male,50-54	0.737	8.35	0.6754	0.7986	0.7986	++
Male,55-59	0.863	16.41	0.7214	1.0046	0.7214	- -
Male,60-64	1.028	28.73	0.7327	1.3233	1.1182	+
Female,21-24	0.433	28.95	0.3077	0.5583	0.5243	+
Female,25-29	0.548	4.73	0.5221	0.5739	0.5221	- -
Female,30-34	0.656	29.12	0.4650	0.8470	0.8470	++
Female,35-39	0.76	28.72	0.5418	0.9782	0.9782	++
Female,40-44	0.839	14.56	0.7168	0.9612	0.7168	- -
Female,45-49	0.878	24.01	0.6672	1.0888	0.6704	-/- -
Female,50-54	1.013	4.26	0.9699	1.0561	1.0561	++
Female,55-59	1.054	12.65	0.9206	1.1874	1.1874	++
Female,60-64	1.156	27.47	0.8384	1.4736	0.9859	-
Male, 2-4	0.283	23.77	0.2157	0.3503	0.3503	++
Male, 5-9	0.196	28.78	0.1396	0.2524	0.2524	++
Male, 10-14	0.246	19.67	0.1976	0.2944	0.2942	+/++
Male, 15-20	0.336	1.07	0.3324	0.3396	0.3396	++
Female, 2-4	0.233	25.47	0.1736	0.2924	0.2151	-
Female, 5-9	0.165	28.02	0.1188	0.2112	0.1188	- -
Female, 10-14	0.223	20.36	0.1776	0.2684	0.2105	-
Female, 15-20	0.379	22.73	0.2928	0.4652	0.3604	-
Asthma	1.098	22.29	0.8532	1.3428	1.0851	-
Acute Appendicitis	0.3	11.77	0.2647	0.3353	0.2656	-/- -
Diabetes	1.331	19.66	1.0693	1.5927	1.0693	- -
Congestive Heart Failure	3.79	5.14	3.5954	3.9846	3.9846	++
HIV	5.485	21.18	4.3232	6.6468	4.8978	-
Mental Illness	1.5	0.95	1.4857	1.5143	1.4857	- -

The codes for column "Change" are: (1) "-": robust weight is smaller than nominal weight; (2) "- -": robust weight reaches the lower bound; (3) "-/- -": robust weight is very close to lower bound, due to numerical issues; (4) "+": robust weight is larger than nominal weight; (5) "++": robust weight reaches the upper bound; (6) "+/++": robust weight is very close to lower bound, due to numerical issues.

Table 2.9: Nominal risk scores vs. Robust risk scores

	Insurer 1	Insurer 2	Insurer 3	Insurer 4	Insurer 5
Nominal Risk Score	0.9990379	0.9993510	1.0005282	1.0008137	0.9995409
Robust Risk Score	0.9990487	0.9992846	1.0004502	1.0007771	0.9995645
Change in Risk Score	0.0011%	-0.0066%	-0.0078%	-0.0037%	0.0024%
Nominal money transfer	-192413.36	-129803.99	105639.84	162732.18	-91822.65
Robust money transfer	-190265.94	-143077.69	90036.45	155424.44	-87108.02
Change in money transfer	-1.1160%	10.2260%	-14.7704%	-4.4907%	-5.1345%
	Insurer 6	Insurer 7	Insurer 8	Insurer 9	Insurer 10
Nominal Risk Score	1.0007512	1.0007304	1.0003757	1.0000186	0.9988525
Robust Risk Score	1.0007750	1.0006677	1.0004926	1.0000597	0.9988801
Change in Risk Score	0.0024%	-0.0063%	0.0117%	0.0041%	0.0028%
Nominal money transfer	150238.92	146073.36	75135.25	3714.90	-229494.45
Robust money transfer	154993.01	133534.91	98516.48	11931.70	-223985.34
Change in money transfer	3.1644%	-8.5837%	31.1189%	221.1848%	-2.4005%

We can see from Table 2.9 that although the percentage changes in risk scores are small, the changes in actual money transfers are significant. The reason is that the relative change in risk score is calculated as $\frac{RS - RSN}{RSN}$, while the relative change in actual money transfer is calculated as $\frac{(RS-1) \times N \times C - (RSN-1) \times N \times C}{(RSN-1) \times N \times C}$, or equivalently $\frac{RS - RSN}{RSN - 1}$: the numerator stays the same but the denominator is not and this can create significant changes because the risk scores are close to 1 to begin with. Note that RS, RSN, N, and C denote robust risk score, nominal risk score, number of enrollees in the plan, and capitation payment per person, respectively. In the example above, 4 out of 10 payers observe a relative change in actual money transfer higher than 10%.

2.8 Conclusions

In this chapter, we have investigated how to mitigate the impact of parameter uncertainty on risk scoring in healthcare. An example related to hospital ranking using adjustment factors was provided to demonstrate the need for robustness. We pre-

sented an approach to compute robust risk scores. Our methodology involves solving a series of linear programming problems and thus is highly tractable. Future work includes using budget uncertainty and addressing uncertainty related to the health status of previously uninsured customers entering the system due to the Affordable Care Act of 2010.

Chapter 3

Plan Efficiency Evaluation in Health Insurance Exchanges

3.1 Introduction

The establishment of the Health Insurance Marketplace through the Affordable Care Act has opened an important channel for individuals to obtain health coverage [3]. It allows individuals to compare different health plans and shop health insurance on a web-based portal. The exchanges are created and regulated by the federal or state governments. While the wide range of plans available on exchanges during the 2015 open enrollment period suggests a robust health insurance market, the seminal work by Iyengar and Lepper [55] suggests that extensive choice can be overwhelming for customers and lead to greater procrastination before making a choice and/or dissatisfaction afterward. Specifically, Iyengar and Lepper [55] claim that although psychological benefits of provision of choice do exist, people are more satisfied when they are given a more limited number of options, which they can compare more

easily in order to reach an informed decision. In the context of health insurance exchanges, Day and Nadash [31] advise that states should align themselves on Massachusetts in offering only a reasonable but limited set of choices. Similarly, Saxena and Holobinko [67] demonstrate that too much choice can inhibit consumers and reduce the satisfaction they derive from their decision. They also argue that consumers are highly price sensitive when shopping on exchanges.

The objective of this chapter is to provide an analytical methodology to help individuals narrow down plan choices by identifying plans that are dominated by competitors' offerings in terms of premium, metal level, maximum out of pocket payment and plan type (such as HMO, POS, EPO or PPO). If a customer makes her plan choice based on the features above, it is not in her best interest to select a dominated plan, but a customer who analyzes alone the many plans available on the exchanges may not realize it and thus may make a less-than-optimal decision for herself. Our approach also provides payers with a novel way to analyze their own plans in the HIX landscape. The high-level goal of this methodology is to contribute to a wider discussion on making plans offered on the exchanges as compelling as possible for health exchange customers. We illustrate our methodology using a federal-run exchange (Pennsylvania) and a state-run exchange (Massachusetts).

3.2 Methodology and Data Sources

The following factors should be considered when comparing marketplace plans: (a) Plan metal level: Catastrophic (for individuals under the age of 30 with hardship exemptions) Bronze, Silver, Gold, Platinum, in increasing order of actuarial value, (b) Monthly premium, (c) Out-of-pocket costs - including deductible, copayments, co-insurance - up to the out-of-pocket maximum, (d) Type of insurance plan and

provider network: HMO, POS, EPO or PPO, because each type has different limitations on usage of in-network versus out-of-network doctors and referrals for specialty care, and (e) Benefits: although all marketplace plans have the same essential health benefits, cover pre-existing conditions and offer free preventive services, some plans offer additional benefits.

Intuitively, a plan that has less choice (is of a "worse" type, offering less choice) and higher premium than another plan should have a better metal level or a lower maximum out of pocket to justify the higher premium. If it does not, the plan is said to be, in our terminology, "inefficient" or "dominated." We do not compare plans issued by the same payer because we assume that those plans are priced in a consistent fashion, i.e., a single decision-maker prices plans rationally, with a more desirable plan always priced higher.

Our objective is to identify "dominated" plans in this framework and quantify their excess premium, i.e., the amount by which their premium should be decreased to be made efficient in this framework or the amount that the payer must take great care in justifying to convince potential enrollees' of the plan's value proposition. We quantify the excess premium in two different ways: by re-running a regression with, as explanatory variables, either plans by non-dominated payers that are assigned a good letter grade, such as B or above (Method 1), or plans that are rarely dominated by other plans, with the precise threshold percentage set by the analyst (Method 2).

Our **methodology** consists in the following four steps:

1. We identify inefficient plans using an algorithm that - for plans issued by different payers - compares types (HMO, POS, EPO or PPO), premiums, metal levels and maximum out of pockets.

2. We run a regression explaining plan premiums using plan features including pay-

er's name, and check whether payer's name is a significant variable. If it is, we assign a letter grade to each payer using Step 3. Otherwise, we go to Step 4.

3. (**Method 1 only**) (Step a) For each payer, we compute the percentage of non-dominated plans offered on the exchange. We cluster payers of similar percentages and assign each cluster a letter grade. (Step b) We rerun the regression of Step 1 with the payers graded B or higher only. We then compute the estimated premium that a plan by a payer ranked strictly below B should have had in this model and compare it to the premiums of a plan with the same attributes offered by a payer ranked B or higher. The excess premium for a plan by a payer graded strictly below B is the smallest difference between its plan's premium and the premium of a similar plan offered by a high-graded payer, when positive.

4. (**Method 2 only**) Re-run the regression using only plans that are rarely dominated by other plans, with the threshold percentage being set by the analyst. Then, compute what the premiums of dominated plans should have been based on their features for them to be non-dominated. The decrease in premium that the payer should consider - or the part of the premium that he should more clearly justify - is the smallest difference between the premium that a non-dominated or efficient plan should have given the plan features and the regression results.

The **algorithm** to select efficient plans is described as follows. When selecting efficient plans, we take four factors into consideration: premium, level, type, and out-of-pocket maximum. We assume throughout that plans from the same company are priced appropriately based on their characteristics, i.e., there is no inefficient plan within the pool offered by a given company. Hence, we only compare plans from different payers. Our goal is to identify those plans that have higher premiums with worse benefits. We label a plan P as "inefficient" when we can find another plan Q such that P's type (such as HMO or PPO) is the same or worse than Q's, P's

premium is higher than Q's, and either P's metal level is worse than Q's or its metal level is the same but its out-of-pocket maximum is higher than or equal to Q's. If no such Plan Q can be found for a given plan P, Plan P is said to be "efficient". We perform those checks using the MATLAB computer programming software.

To investigate whether some payers are more efficient (offer more efficient plans) than others, we rank payers according to the percentage of efficient plans in their HIX portfolio. The higher the percentage is, the less likely it is that a payer's plans are dominated. Based on those percentages, we use a three-level grading scale: A, B and C, to group payers, when appropriate (i.e., when payer's name is a significant variable). The ranks and grades for insurers in PA and MA are shown in Table 3.6 and Table 3.8, respectively.

For the purpose of plan evaluation, we use linear regression to determine the weights of attributes of plans in the marketplace. The dependent variable is premium. The independent variables include numeric variables: deductible, out-of-pocket maximum; categorical (dummy) variables: company, metal level (bronze, etc.), type (HMO, etc.). We also run individual regression models for each plan level - bronze, silver, gold ad platinum - using stepwise variable selection.

3.3 Results

3.3.1 Data Visualization in the Pennsylvania HIX

Figure 3.1 shows the distribution of plans based on metal level, premium and deductible.

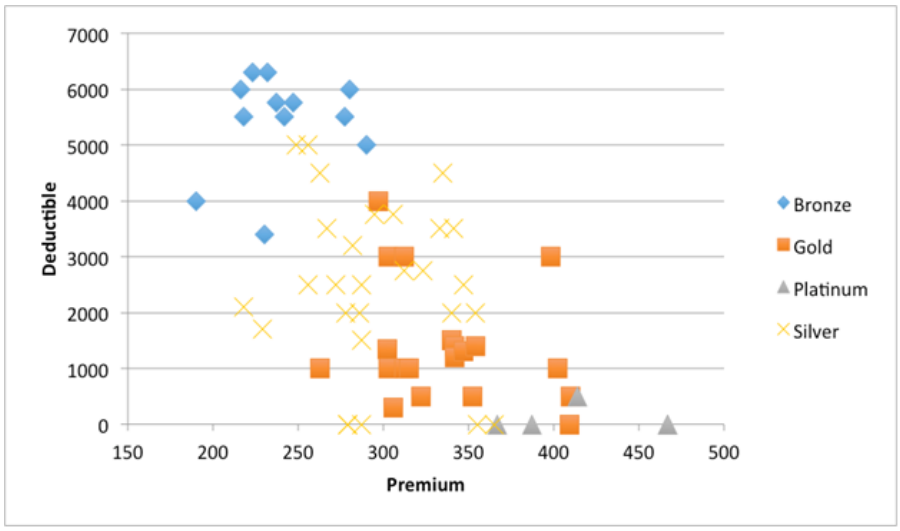


Figure 3.1: Plans in a (premium, deductible) plane according to metal level, PA HIX

We notice that the two plans pointed by arrows have very similar premiums (230 vs 229). However, they have quite different deductibles: the bronze plan is $3400, and the silver plan is $1700. The bronze plan seems to be inefficient (dominated by the other plan) because bronze plans are worse than silver plans in terms of cost sharing, and the bronze plan even has a much higher deductible. It is not the end of the story yet. We take a closer look at these two plans. The bronze plan is the Highmark Health Savings Blue PPO 3400, and the silver plan is the Highmark Flex Blue PPO 1700. They are both PPO plans from Highmark. The only difference is that the bronze plan is HSA and the silver plan is FSA, where HSA has more flexibility than FSA. The value of HSA contributes to the total value of the bronze plan and possibly explains the same premium of these two plans. Another example is as follows. In this case, although the second plan has zero deductible, its other benefits such as doctor visits and ER visits are worse than the first plan. This might lead to their similar premiums.

Table 3.1: Plans from the same company - Example

Company	Plan	Level	Type	Premium	Deductible	OOP max	Primary doc	Specialist doc	ER	Generic drugs
CBC	Healthy Benefits PPO 2000.0	Silver	PPO	$354	$2,000	$6,350	$30	$50	$250	$20 co-pay AD
CBC	Healthy Benefits PPO 0.50	Silver	PPO	$355	$0	$6,350	$50	50%	50%	50%

Therefore, we assume that plans from the same company are priced appropriately based on their characteristics, aka there must be a reason why two plans from the same company have the same premium but quite different deductibles. However, if we compare plans from different companies we would notice that some plans are indeed inefficient. For example, we look at the following two plans:

Table 3.2: Plans from different companies - Example 1

Company	Plan	Level	Type	Premium	Deductible	OOP max	Primary doc	Specialist doc	ER	Generic drugs
Assurant	Health-Bronze Plan 002	Bronze	PPO	$290	$5,000	$6,350	$35	$35	$100 copay BD/25% coins AD	25% coins
Geisinger	Choice-Marketplace PPO 30/50/5000	Silver	PPO	$256	$5,000	$6,000	$30	$50	$250	$3 AD

Obviously, the Geisinger plan has better benefits with lower premium. In this case, we could say that the Assurant plan is inefficient. Another example is as follows. In this case, all three plans are silver plans. The Highmark plan has best benefits thus has highest premium. However, the Keystone plan seems to be inefficient due to its less flexibility (HMO vs. POS) and higher OOP max.

Table 3.3: Plans from different companies - Example 2

Company	Plan	Level	Type	Premium	Deductible	OOP max	Primary doc	Specialist doc	ER	Generic drugs
Geisinger	Marketplace POS 25/50/2500	Silver	POS	$256	$2,500	$5,500	$25	$50	$250	$3
Highmark	Health Savings Blue PPO 2500	Silver	PPO	$287	$2,500	$3,500	10% coins	10% coins	10% coins	10% coins
Keystone	Healthy Benefits Value HMO 2500.0	Silver	HMO	$272	$2,500	$6,350	$20	$40	$300	$20 AD

3.3.2 Plan Efficiency Evaluation Results in the Pennsylvania HIX

We run a regression model both with and without stepwise variable selection, to remove non-significant variables. Our next step is to include all Company variables, Level variables, and Deductible then rerun the regression model. Among all the independent variables left in the model, all have p-values smaller than 0.05 (significant) except the dummy variable referring to the Keystone payer and Deductible (shown in Table 3.4). A summary of the regression models, including R-square values and variables included, is presented in Table 3.5. The QQ plot is approximately a straight line, and the residuals are Normal. We therefore argue that this linear model is valid.

Table 3.4: Regression results with stepwise selection - PA

Variable	Label	DF	Parameter estimate	Standard error	t value	Pr> $\|t\|$
Intercept	Intercept	1	231.8229	11.78069	19.68	<.0001
Company_new2	Geisinger	1	-15.2616	7.25866	-2.1	0.04
Company_new3	Keystone	1	-4.52886	6.76242	-0.67	0.5058
Company_new4	Aetna	1	28.78683	7.44184	3.87	0.0003
Company_new5	Assurant	1	70.08725	8.50878	8.24	<.0001
Company_new6	CBC	1	74.45939	6.92121	10.76	<.0001
Level_new2	Silver	1	49.5767	8.38377	5.91	<.0001
Level_new3	Gold	1	97.71053	10.35155	9.44	<.0001
Level_new4	Platinum	1	163.9487	14.0219	11.69	<.0001
Deductible	Deductible	1	-0.00323	0.00199	-1.63	0.1097

Next, we look at the individual regression models for each metal level. Since there are few plan observations at the bronze and platinum levels, we only investigate the cases of the silver and gold plans. The regression results illustrate that the identity of the payer is the most important attribute driving plan premiums. Results of all regression models on the PA data (for the 18015 zip code and an adult in her late thirties) are summarized in Table 3.5.

Table 3.5: Summary of regression models - PA

Model	No. of observ.	R-square	Variables
All variables	66	0.9273	all
Stepwise selection	66	0.9263	all companies, all levels, deductible
Bronze only	12	0.8109	Assurant, CBC, OOP_Max
Silver only	28	0.8501	Aetna, Assurant, CBC, Deductible
Gold only	21	0.7816	Geisinger, Assurant, CBC
Platinum only	5	1	Geisinger, HMO, Deductible

We apply the algorithm to identify efficient plans to PA data. Key results are summarized in Table 3.6. We then rank the payers based on their efficiency ratios. Here, Highmark and Geisinger have the highest percentage of absolutely efficient plans and thus, according to our framework, should be recommended to individuals seeking coverage on the health insurance exchanges.

Table 3.6: Summary of inefficient plans & ranks of companies - PA

	Aetna	Assurant	CBC	Geisinger	Highmark	Keystone
Company sum of dominated times	85	38	109	8	6	64
Number of plans in each company	10	6	11	11	16	12
Number of comparisons	560	360	605	605	800	648
Percent of dominated by others	15.18%	10.56%	18.02%	1.32%	0.75%	9.88%
Number of absolutely efficient plans	0	0	0	3	10	0
Percent of absolutely efficient plans	0.00%	0.00%	0.00%	27.27%	62.50%	0.00%
Rank	5	4	6	2	1	3
Grade	C	B	C	A	A	B

It is also interesting to look at the premiums from the perspective of PA general acute care hospitals' total margin. Figure 3.2 below is the average total margins over

regions in PA for fiscal year 2013. Figure 3.3 is the coverage areas served by blue plans in PA.

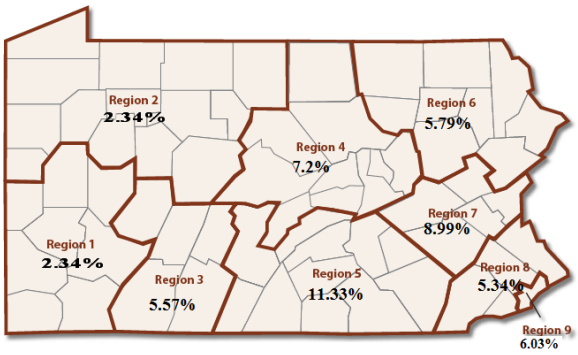


Figure 3.2: Hospital total margin - PA

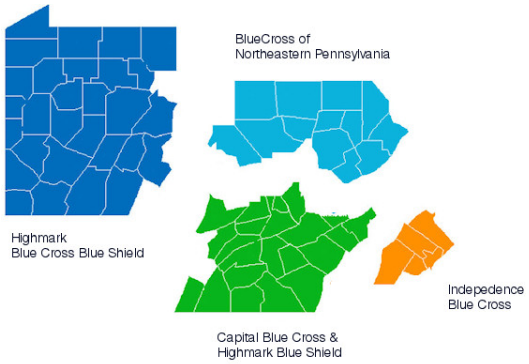


Figure 3.3: Coverage area - PA

source: http://pablueagent.com/coverage.php

Hospitals in region 5 and region 7 have the highest margins, while these two regions are areas where CBC mostly cover. This observation might indicate one of the reasons why CBC generally has higher premiums and more number of inefficient plans: CBC pays more to hospitals than other insurers, thus they need to increase premiums to keep in business, which might leads them to less competitive position in the market. On the contrary, hospitals in areas covered by Highmark in general have very low margins, which supports our conclusion from another perspective.

3.3.3 Plan Efficiency Evaluation Results in Massachusetts HIX

We use data obtained for the 02142 zipcode for an adult in her late thirties. Only one plan offered is PPO, and seven plans are EPO; however, these seven EPO plans are all from United Health Care, and all United Health Care plans are EPO plans. As a result, we exclude "Type" in the regression model since it is already captured

in the "Company" ("Payer") variable.

Compared to Pennsylvania plans, Massachusetts plans have far fewer deductible levels (see [89] for the MA counterpart to Figure 3.1). We run a regression model to identify the factors attributed to plan premiums. The R-square of this model is 0.8402, which is very good. All variables except the payer Tufts and Out-of-pocket maximum are significant at the 5% level. The residual plots (not shown here) suggest that this model fits quite well. Specifically, the QQ plot is almost a straight line, and the residuals follow a Normal distribution. We also apply stepwise variable selection to MA plans but all variables are selected by the program, leading back to the original model.

We also apply regression models to plans at each metal level. More plans are offered in the 02142 Massachusetts zip code than in the 18015 Pennsylvania zip code, so the results are more convincing. Also note that the MA plans studied here tend to have higher metal levels, where silver is the dominant metal level in PA and gold is dominant in MA. The regression results shown in Table 3.7 indicate that the payer is still the most important attribute in plan premiums. Therefore, it is meaningful to compare plans from different companies and rank companies based on plan efficiency.

Table 3.7: Summary of regression models - MA

Model	No. of observ.	R-squared	Variables
Model with all variables	106	0.8402	all
Stepwise selection	106	0.8402	all
Bronze only	14	0.4675	MA_BCBS, OOP_Max
Silver only	20	0.6077	Fallon, Harvard, MA_BCBS, United
Gold only	48	0.6803	Fallon, Harvard, MA_BCBS, Tufts, United
Platinum only	24	0.7408	Fallon, Harvard, MA_BCBS, Tufts, United

We apply the algorithm to all MA plans. Key results are summarized in Table 3.8.

Companies in MA are ranked and given grades in the same fashion as companies in PA. Based on this framework, plans provided by United HealthCare, Boston Medical Center and Minuteman Health particularly deserve recommendation.

Table 3.8: Summary of inefficient plans & ranks of companies - MA

	Ambetter	BMC	Fallon	Harvard	BCBS	Minuteman	NHP	Tufts	United
Company sum of dominated times	71	9	602	231	243	14	39	90	2
Number of plans in each company	9	7	27	13	10	9	9	13	7
Number of comparisons	855	679	2079	1183	940	855	855	1183	679
Percent of dominated by others (%)	8.3	1.33	28.96	19.53	25.85	1.64	4.56	7.61	0.29
Number of absolutely efficient plans	1	3	0	1	1	4	1	2	5
Percent of absolutely efficient plans (%)	11.11	42.86	0	7.69	10	44.44	11.11	15.38	71.43
Rank	6	2	9	7	8	3	4	5	1
Grade	B	A	C	C	C	A	B	B	A

3.4 Recommendations for Other States and Health Policy Makers

To extend our methodology to other states besides MA and PA, analysts should first identify any unique characteristics of plans in the state once they have obtained the summary statistics of plan premium and benefits. For example, the types of MA plans (HMO, PPO and EPO) are different from the PA plans (HMO, PPO and POS), and it turns out that we can discard the Type variable in the MA regression model. The second step is to utilize regression models to decide whether a payer (name) is the most important indicator of inefficiency/efficiency. If it is, then we feel justified in assigning grades to payers; otherwise, we have to implement a different method (Method 2) to compute premiums.

Health payers with offerings less compelling (as quantified by this framework) than their competitors may wish to discuss internally the assumptions that led their

analysts to pricing the plan higher than what the regression model recommends, potential ways to realize savings in order to decrease premiums, or a stronger positioning of their plan to justify the apparent inefficiency using a different quantitative metric.

We now illustrate possible recommendations to health insurance companies using the Capital Blue Cross (CBC) payer in Pennsylvania. CBC has a total of eleven plans but zero absolutely efficient plan, and is thus given the grade of C. In order to give suggestions to CBC regarding the size in premium reduction that would make its plans more competitive in our model, we first conduct regression on plans from payers graded A or B (Highmark, Geisinger, Keystone, and Assurant) since they are the "aspirational peers" of payers ranked C. Then we apply the regression results to CBC plans in order to get approximate premiums for plans with such characteristics offered by A or B payers. Based on 45 plans from grade A and grade B companies, the R-squared of the model is 0.8919 and the adjusted R-squared is 0.8602, both of which are very high. Although the F statistic is significant, some of the variables are not very significant at the individual level. As a result, we apply stepwise variable selection on the model. After the stepwise variable selection, the Type variables and Deductible are excluded from the model. We include the selected variables and run the regression model again. The R-square drops down slightly to 0.8896 but the adjusted R-square goes up to 0.8687.

Our next step is to apply the regression coefficients to CBC plans. Table 3.9 shows the estimated premiums of each CBC plan if they were plans from grade A and B companies. (Those estimated premiums are all smaller than the actual ones.) Well-priced plans are in bold. This gives CBC an estimate of the premium decreases they should achieve in order to stay competitive in the market. Alternatively, this is a measure of how much CBC feels they will attract sicker patients due to their name

recognition.

Table 3.9: Estimated premiums - CBC Healthy Benefits (HB), PA, Method 1.

Plan	Level	Type	Deduct.	OOP Max	Prem.	Est_Pre_Highm	Est_Pre_ Geis.	Est_Pre_Keyst.	Est_Pre_Assurant
PPO 5500.50	**Bronze**	**PPO**	**5500**	**6350**	**277**	**220.86**	**202.67**	**220.73**	**289.44**
PPO 3000.0	Gold	PPO	3000	6350	398	321.54	303.35	321.42	390.12
PPO 1000.0	Gold	PPO	1000	6350	402	321.54	303.35	321.42	390.12
PPO 500.0	Gold	PPO	500	6350	410	321.54	303.35	321.42	390.12
PPO 0.0.10	Platinum	PPO	0	6350	467	389.63	371.44	389.51	458.21
PPO 4500.50	**Silver**	**PPO**	**4500**	**6350**	**335**	**269.29**	**251.11**	**269.17**	**337.87**
PPO 3500.0	Silver	PPO	3500	6350	341	269.29	251.11	269.17	337.87
PPO 2500.0	Silver	PPO	2500	6350	347	269.29	251.11	269.17	337.87
PPO 2000.0	Silver	PPO	2000	6350	354	269.29	251.11	269.17	337.87
PPO 0.50	Silver	PPO	0	6350	355	269.29	251.11	269.17	337.87
PPO 0.0	Silver	PPO	0	6350	365	269.29	251.11	269.17	337.87

The estimated premiums give us ranges of premiums for efficient plans with same benefits. Out of the eleven plans, two CBC plans' premiums (highlighted in bold) are within the ranges determined by efficient plans, solely because they have lower premiums than Assurant's corresponding plans. The reason could be that Assurant is only slightly better than CBC (rank 4th vs. 6th) in spite of being ranked B instead of C, so that occasionally CBC plans are more efficient than Assurant's.

We calculated the estimated premiums for Aetna plans (grade C) in a similar fashion (shown in Table 3.10). We notice that Aetna plans are always more expensive than plans from Highmark, Geisinger and Keystone with the same benefits. However, they are always cheaper than the corresponding Assurant plans as well. This is not surprising because their ranks are even closer (rank 4th vs. 5th). As a result, it is more likely that Aetna has more efficient than Assurant compared to CBC.

Table 3.10: Estimated premiums - Aetna

Plan	Level	Type	Deduct.	OOP Max	Prem.	P_Highm.	P_Geis.	P_Keyst.	P_Assu.
Deduct. only HSA eligible HMP	Bronze	HMO	6300	6300	223	220.97	202.78	220.84	289.54
Deduct. only HSA eligible OAMC	Bronze	POS	6300	6300	232	220.97	202.78	220.84	289.54
$20 copay OAMC	Bronze	POS	5750	6600	247	220.33	202.15	220.21	288.91
$20 copay HMO	Bronze	HMO	5750	6600	237	220.33	202.15	220.21	288.91
$5 copay HMO	Gold	HMO	1400	5000	342	324.39	306.20	324.27	392.97
$5 copay OAMC	Gold	POS	1400	5000	354	324.39	306.20	324.27	392.97
$10 copay HMO	Silver	HMO	3750	6600	295	268.77	250.58	268.64	337.34
$10 copay OAMC	Silver	POS	3750	6600	306	268.77	250.58	268.64	337.34
$5 copay 2750 HMO	Silver	HMO	2750	6000	312	270.03	251.84	269.91	338.61
$5 copay 2750 OAMC	Silver	POS	2750	6000	323	270.03	251.84	269.91	338.61

We also run Method 2, with a threshold percentage of plan being dominated by at most 10% other plans. There are 44 such plans. The R-square is 0.9068. The highest excess premium is $53 (for a monthly premium of $323), achieved for Aetna Silver $5 copay 2750 OAMC. (2750 is the deductible for the plan and OAMC stands for Open-Access Managed Choice.) In fact, the four plans with the highest excess premium are Aetna plans. This may not mean that they are overpriced, but Aetna should make sure that its plans' value proposition is well understood. The most underpriced plan is Capital Blue Cross Healthy Benefits PPO 5500.50, which has a monthly premium of $277 although the model estimates the premium at $296, and the second most underpriced plan is Keystone Healthy Benefits Value HMO 1000.0 (monthly premium $315, estimated premium $325). Those results are shown in Table 3.11.

Table 3.11: Estimated premiums - PA, Method 2

Plan	Level	Type	Deductible	OOP_Max	Prem.	Est_Prem.	Decrease
Aetna Gold $5 copay HMO	Gold	HMO	1400	5000	342	325	17
Aetna Gold $5 copay OAMC	Gold	POS	1400	5000	354	325	29
Aetna Silver $10 copay HMO	Silver	HMO	3750	6600	295	270	25
Aetna Silver $10 copay OAMC	Silver	POS	3750	6600	306	270	36
Aetna Silver $5 copay 2750 HMO	Silver	HMO	2750	6000	312	270	42
Aetna Silver $5 copay 2750 OAMC	Silver	POS	2750	6000	323	270	53
Assurant Health - Bronze Plan 002	Bronze	PPO	5000	6350	290	289	1
Assurant Health - Gold Plan 002	Gold	PPO	0	6350	409	387	22
Assurant Health - Silver Plan 002	Silver	PPO	2000	6350	340	331	9
Healthy Benefits (HB) PPO 5500.50	Bronze	PPO	5500	6350	277	296	-19
Healthy Benefits (HB) PPO 3000.0	Gold	PPO	3000	6350	398	394	4
Healthy Benefits (HB) PPO 1000.0	Gold	PPO	1000	6350	402	394	8
Healthy Benefits (HB) PPO 500.0	Gold	PPO	500	6350	410	394	16
Healthy Benefits (HB) PPO 4500.50	Silver	PPO	4500	6350	335	339	-4
Healthy Benefits (HB) PPO 3500.0	Silver	PPO	3500	6350	341	339	2
Healthy Benefits (HB) PPO 2500.0	Silver	PPO	2500	6350	347	339	8
Healthy Benefits (HB) PPO 2000.0	Silver	PPO	2000	6350	354	339	15
Healthy Benefits (HB) PPO 0.50	Silver	PPO	0	6350	355	339	16
Healthy Benefits (HB) PPO 0.0	Silver	PPO	0	6350	365	339	26
Keystone Healthy Benefits Value HMO 1000.0	Gold	HMO	1000	6350	315	325	-10
Keystone Blue Cross Value HMO 500.0	Gold	HMO	500	6350	322	325	-3
Keystone Healthy Benefits Value HMO 0.0	Silver	HMO	0	6350	287	270	17

There is concern that some insurance companies might set their prices lower than others on purpose to attract new customers in the exchanges, and raise their prices in later years. Since the exchanges are new to both insurers and customers, prices can be

adjusted over time to reflect learning from new information. Although it is possible that some insurers appear to be more efficient, while some others being "inefficient" due to short-term pricing strategy, our approach still provides a novel way to aid customer's choice of plan and to help insurers analyze their plans. More years of data could be incorporated when the Health Insurance Marketplaces become more mature, and the long-term efficiency of plans could be evaluated in our framework.

3.5 Conclusions

This chapter investigates the important drivers of plan premiums in the health exchanges, with examples drawn from the Pennsylvania and Massachusetts health insurance marketplaces. We propose a simple algorithm to identify plans that are dominated by competitors' offerings in terms of premium, metal level, maximum out of pocket payment and plan type. We assign letter grades to payers, and use linear regression results to quantify the amount by which the premium of a dominated plan should be reduced to make it competitive in our framework. This chapter provides important quantitative tools to guide the discussions between payers and policy-makers regarding HIX. Our approach also provides payers with a novel way to analyze their own plans in the HIX landscape.

Chapter 4

Analysis of Medicare Prescription Drug Coverage Enrollment

4.1 Introduction

The Medicare Prescription Drug Improvement and Modernization Act of 2003 provides outpatient prescription drug coverage for Medicare beneficiaries through private insurers. The coverage is available for Medicare beneficiaries in two ways: Medicare Advantage prescription plans (MAPDs), and stand-alone prescription plans (PDPs). This chapter provides an overview of these two types of prescription drug plans and identify factors associated with Medicare beneficiaries' choices over MAPDs and PDPs in all counties in the United States using beta regression.

4.2 Background: PDPs vs. MA-PDs

Medicare is a national social insurance program administered by the U.S. federal government since 1966. It covers people aged 65 and over and people with permanent disabilities. Currently about 54 million people are under coverage of Medicare. It has four parts: (1)Part A: Hospital Insurance, which covers most medically necessary hospital, skilled nursing facility, nursing home and hospice care; (2)Part B: Medical Insurance, which covers most medically necessary doctors' services, preventive care, durable medical equipment, hospital outpatient services, laboratory tests, x-rays, mental health care, and some home health and ambulance services; (3)Part C: Medicare Advantage plans, which allows private health insurance companies to provide Medicare benefits since the passage of the Balanced Budget Act of 1997; (4)Part D: Prescription Drug Plans, which provides outpatient prescription drug coverage. Among these four parts, Part D was established by the Medicare Modernization Act of 2003 (MMA) and went into effect in 2006.

Medicare beneficiaries have two ways for receiving prescription drug coverage starting in 2006. One way is to enroll in the Medicare Advantage plans (mainly HMOs and PPOs) then enroll in their Part D prescription drug plans (MAPD). Medicare Advantage plans are provided by private insurance companies that have a contract with Medicare. In 2014, 83% of Medicare Advantage plans offer prescription drug coverage, and 50% provide some coverage in what is known as the "coverage gap", after a spending limit is exceeded but before the maximum out-of-pocket payment is attained. All MAPD enrollees receive a 50% discount on brand-name drugs in the gap, beginning in 2011. Since 2011, all Medicare Advantage plans have been required to limit beneficiaries' out-of-pocket spending to no more than $6,700. Medicare Advantage beneficiaries need to enroll in both Part A and Part B. Many MA-PDs

provide prescription drug coverage and other supplemental benefits available for no additional premium. However, Medicare Advantage plan monthly premiums (if any) are in addition to Medicare A (if required) and Medicare B premiums.

The other way to get prescription drug coverage is to enroll in Prescription Drug Plans (PDP) provided by Medicare directly. In order to enroll into a Medicare Part D PDP, enrollment in either Medicare Part A and/or Medicare Part B is required. Medicare Part D premiums are in addition to Medicare Part A (if any) and/or Part B premiums; however, original Medicare doesn't have the cap on out-of-pocket spending.

In 2014, 37 million out of 54 million Medicare beneficiaries enrolled in part D plans. Among these part D plan enrollees, more than 37% of all Medicare beneficiaries enrolled in part D in late 2013 were in an MA-PD, with the remaining 63 percent in a freestanding PDP. Another large group of Medicare beneficiaries get prescription drug coverage under plans offered by former employers such as the Department of Veterans Affairs. Enrollment in Medicare drug plans is voluntary, with the exception of beneficiaries who are dually eligible for both Medicare and Medicaid and certain other low-income beneficiaries who are automatically enrolled in a PDP if they do not choose a plan on their own.

In terms of plan availability, MAPD plans and PDP plans are quite different. The national average number of PDP plans in 2012 is 31, and the national average number of MAPD plans is 60; however, the number of PDP plans in each state is quite similar as almost every state has about 30 PDP plans; while the number of MAPD plans in each state varies from 1 to 250.

Cline et.al [27] use multivariate probit models to describe factors associated with Medicare beneficiaries' choices in enrolling any part D plans, and their choices of

an MAPD plan given enrollment in the part D program. Their data is collected from surveys of 5000 community-dwelling adults in CMS Region 25, which includes 7 states: Iowa, Minnesota, North Dakota, South Dakota, Nebraska, Montana and Wyoming. Only 32.4% of responses are actually used in the analysis. They find that factors including rurality, plan price, perceived future need for medications, and preferences are driving people's choices of enrolling in a part D plan; while rurality, state of residence, and number of diagnosed medical conditions are contributing to people's decision to enroll in a MAPD plan given enrollment in a part D plan. As their data is collected from surveys and is detailed at member level, they are able to conduct probit model since their response variables would be binary.

However, data at member level is not always available. Therefore, the objective of this chapter is to identify factors associated with penetration rates of MAPD plans and PDP plans based on publicly available census data and plan data. We utilize a relatively new regression method, called beta regression, when describing factors associated with people's choice. This method is particularly suitable for our penetration data sine it incorporates natural properties of variables whose value lies in the open interval (0, 1) such as proportions or penetration rates instead of being binary (0 or 1, yes or no) such as enrollment decision for each member.

4.3 Data Description

Our data combines three sources: penetration data, census data, and part D plan data. All data are from year 2012. The dependent variables are penetration rates of MAPD and PDP, which are retrieved from Kaiser Family Foundation's "MA-PD Plan Enrollment as a Percent of Total Medicare Population" table [42]. The reason why we don't use the Medicare Advantage/Part D Contract and Enrollment Data

available on CMS's website [37] is that CMS's data has the total MA enrollment. However, what we need is the PD enrollment data only, which should be lower than total MA enrollment. We use the PDP penetration data from Kaiser Family Foundation's database to be consistent with our choice of MAPD penetration data. When number of enrollees in one county is 10 or less, the penetration rate is reported as zero due to the privacy laws of HIPAA. We exclude Alaska because Alaska is the only state that has zero MAPD penetration rate, while its PDP penetration rate is 39%. According to Kaiser Family Foundation, there were no private insurance companies in Alaska offering Medicare Advantage plans in 2012. There were 69 beneficiaries who were enrolled in Medicare Advantage plans, which made up less than 0.1% of all the Medicare beneficiaries in the state. Since PDP is absolutely dominant in Alaska, comparison between MAPD and PDP penetration rates would not be meaningful.

Initially, there are 36 independent variables including demographic variables, condition variables, cost variables and plan variables. These data come from three sources: the "State/County Table - All Beneficiaries" from CMS's public use files [39], the "State & County QuickFacts" from US Census Bureau [21], and "Plan & Premium Information for Medicare Plans Offering Part D" from CMS [38]. Most of the variables are from the public use files as they represent Medicare beneficiaries' data, while a few variables come from the census due to significant large number of missing values or lack of data in the public use files. We have data for 3029 counties in the United States. The correlations between independent variables and penetration rates are provided in table 4.1.

Table 4.1: Correlations of variables

	MAPD_rate	*PDP_rate*	*total_rate*
cnt_mapd	0.57	-0.44	0.06
cnt_pdp	-0.06	0.08	0.05
Average_Age	-0.28	0.19	-0.07
Percent_Female	-0.06	0.19	0.19
Percent_Male	0.06	-0.19	-0.19
Percent_Non_Hispanic_White	-0.04	0.08	0.07
Percent_African_American	-0.04	0.10	0.09
Percent_Hispanic	0.08	-0.05	0.04
Percent_Other_Unknown	0.06	-0.13	-0.10
Bachelor_degree_or_higher_perc	0.15	-0.32	-0.26
Homeownership_rate	-0.05	0.07	0.03
Per_capita_money_income	0.09	-0.27	-0.26
Persons_below_poverty_level_per	-0.06	0.20	0.20
Percent_Eligible_for_Medicaid	0.07	0.17	0.31
Average_HCC_Score	0.21	-0.06	0.16
Percent_of_heart_attack	-0.06	0.09	0.06
Percent_of_atrial_fibrillation	-0.02	0.05	0.04
Percent_of_kidney_disease	0.22	-0.11	0.11
Percent_of_obstructive_pulmonary	-0.12	0.18	0.10
Percent_of_depression	0.20	-0.05	0.16
Percent_of_diabetes	0.00	0.12	0.16
Percent_of_heart_failure	-0.19	0.26	0.14
Percent_of_ischemic_heart	-0.14	0.17	0.07
Percent_of_breast_cancer	0.07	-0.14	-0.13
Percent_of_colorectal_cancer	-0.17	0.24	0.14
Percent_of_lung_cancer	-0.06	0.05	0.00
Percent_of_prostate_cancer	-0.01	-0.05	-0.09
Percent_of_asthma	0.15	-0.20	-0.11
Percent_of_hypertension	-0.06	0.15	0.13
Percent_of_high_cholesterol	0.10	-0.10	-0.01
Percent_of_arthritis	-0.12	0.15	0.07
Percent_of_osteoporosis	0.07	-0.07	0.00
Percent_of_alzheimer	0.05	0.01	0.09
Percent_of_stroke	0.07	-0.09	-0.04
Part_B_Drugs_Standardized_Costs_perc	0.09	-0.10	-0.04
Percent_of_Beneficiaries_Using_PB	0.14	-0.15	-0.04

All correlations with absolute value higher than 0.20 are highlighted in table 4.1. The number of available MAPD plans is highly correlated with MAPD and PDP penetration rates. The more MAPD plans in a county, the higher their MAPD penetration rate and the lower their PDP penetration rate. Other variables that have relatively high correlations with prescription drug plan penetration rates include

average age of Medicare beneficiaries, percentage of population with Bachelor's degree or higher, per capita income, percentage of Medicare beneficiaries who are also eligible for Medicaid, average HCC score, and a few other medical conditions. Interestingly, most independent variables have opposite correlations with the MAPD penetration rate and with the PDP penetration rate. For example, the correlation between average age and MAPD rate is -0.28, while the correlation between average age and PDP rate is 0.19. This suggests that counties having Medicare population with higher average age tend to see their Medicare members enroll in PDP plans rather than MAPD plans. Also, counties with a higher percentage of people with Bachelor's degree or higher are more likely to enroll in MAPD plans than PDP plans. Hence, these attributes have different impacts on these two types of prescription drug plans.

4.4 Statistical Models on Enrollment Rate of PDP and MA-PD Drug Plans

Our response variables - penetration rates - always lie between 0 and 1. Because of the small range for the values taken by the penetration rates, this suggests that the effect of explanatory variables tends to be non-linear, and the variance tends to decrease when the mean gets closer to one of the boundaries. Linear regression and logistic regression are not attractive for our purposes: linear regression might give us predictions out of the restricted range and logistic regression requires the distribution of response variable to be binomial, while our response variable here is not the results of a set of Bernoulli trials. One possible solution is to transform the response variable then model the transformed response variable on explanatory variables. However, this approach has drawbacks. First, the coefficients cannot be easily interpreted in

terms of the original response variable. Second, asymmetry often exists in rates or proportions, so that inference based on the normality assumption can be misleading.

We introduce another regression model, called beta regression, which is more suitable for our data. Beta regression assumes that the response variable follows a continuous beta distribution and is related to other variables through a regression structure. The estimation of coefficients are conducted by Maximum Likelihood Estimation (MLE). Beta distribution is a family of continuous probability distributions defined on the interval $[0, 1]$ parameterized by two positive shape parameters: mean and precision. The density function of beta distribution is given by

$$\pi(y; p, q) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} y^{p-1}(1-y)^{p-1}, 0 < y < 1 \tag{4.1}$$

where $p > 0$,$q > 0$ and $\Gamma(.)$ is the gamma function. The mean and variance of y are

$$E(y) = \frac{p}{p+q} \tag{4.2}$$

and

$$Var(y) = \frac{pq}{(p+q)^2(p+q+1)} \tag{4.3}$$

respectively. Examples of beta distribution are shown in Figure 4.1.

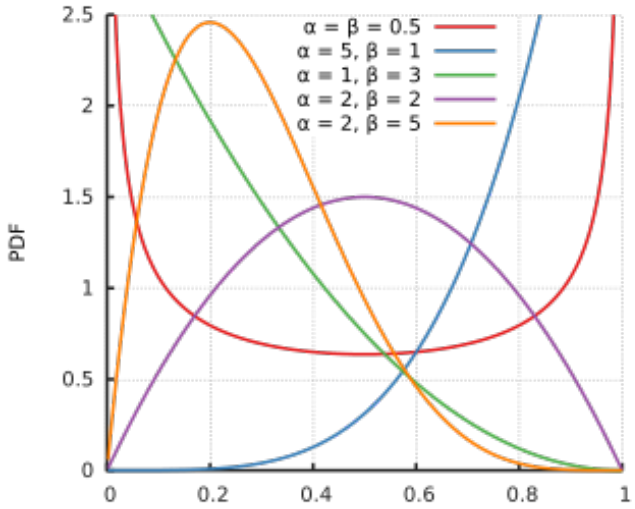


Figure 4.1: Examples of Beta Distribution by Varying Shape Parameters

This conventional parameterization with two shape parameters is not particularly suitable for modeling mean of response variable on explanatory variables. Ferrari and Cribari-Neto [36] propose an alternative parametrization by letting $\mu = \frac{p}{q}$ and $\phi = p + q$ such that μ is the mean of response variable and ϕ is the precision parameter. Then the density function of y becomes

$$f(y;\mu,\phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} y^{\mu\phi-1}(1-y)^{(1-\mu)\phi-1}, 0 < y < 1 \tag{4.4}$$

where $0 < \mu < 1$ and $\phi > 0$.

Now we briefly describe beta regression and how it is different from linear regression of transformed response variable. Let $y_1, ..., y_n$ be independent random response variables, where each y_t for $t = 1, ...n$ follows the beta distribution's density (4.4). The mean of y_t is μ_t and the precision is ϕ. The beta regression model is obtained by assuming that μ_t, which is the mean of y_t that can be written as

$$g(\mu_t) = \sum_{i=1}^{k} x_{ti}\beta_i = \eta_t, \tag{4.5}$$

where $\beta = (\beta_1, ..., \beta_k)^T$ is a vector of unknown regression coefficients, and $x_{t1}, ...x_{tk}$ are observations on k regressors ($k < n$). This indicates the difference between beta regression and linear regression of transformed response variable. In linear regression model would be $g(y_t) = \sum_{i=1}^{k} x_{ti}\beta_i$, while in beta regression it is $g(\mu_t)$ that has a regression structure. Therefore, the regression parameters are interpretable in terms of the mean of response variable in beta regression, and the model is naturally heteroscedastic and easily accommodates asymmetries.

In addition, $g(.)$ is a strictly monotonic and twice differentiable link function that maps (0,1) to $\mathbb{R}$. Possible choices of link function $g(\mu_t)$ include:

- logit function: $g(\mu) = log\frac{\mu}{1-\mu}$,
- probit function: $g(\mu) = \Phi^{-1}(\mu)$, where $\Phi(.)$ is the cumulative distribution function of a standard normal distributed variable,
- complementary log-log link: $g(\mu) = log\{-log(1-\mu)\}$,
- log-log link: $g(\mu) = -log\{-log(\mu)\}$.

One commonly used link function, which is the default one used in "betareg" package in R, is the logit function. In this case, the link function is presented as

$$g(\mu) = log\frac{\mu}{1-\mu} = x_t^T\beta, \tag{4.6}$$

which, after simple manipulation, can be written as

$$\mu_t = \frac{e^{x_t^T\beta}}{1+e^{x_t^T\beta}}, \tag{4.7}$$

where $x_t^T = (x_{t1}, ..., x_{tk}), t = 1, ..., n$. Therefore, the mean of y_t can be easily interpreted by the regression parameters. If the value of ith independent variable is

increased by c units, and all other independent variables remain unchanged. Let μ' be the mean of y under new values of variables, then we have

$$e^{c\beta_i} = \frac{\mu'/(1-\mu')}{\mu/(1-\mu)}. \tag{4.8}$$

It means $e^{c\beta_i}$ equals the ratio of odds under two sets of values. Beta regression uses maximum likelihood estimation to obtain β and ϕ, and it does not have a closed-form solution. Therefore, the parameter estimates are obtained by numerically maximizing the log-likelihood function using a nonlinear optimization algorithm, such as Newton method or quasi-Newton algorithm. To start the algorithm, β estimates from OLS linear regression of transformed response variable can be used as initial values [36].

SAS can implement beta regression using procedures NLMIXED, NLIN or GLIMMIX. A macro called Beta_Regression in SAS is available for implementation purpose [79]. Beta regression can also be implemented in R using package "betareg" [28]. The implementation results using the "betareg" package in R are shown in the next section.

4.5 Beta Regression Implementation Results in R

One limitation of beta regression is that it excludes values 0 and 1. However, due to the privacy law of HIPAA, there are 171 penetration rates that are reported as zero as the numbers of enrollees in those counties are ten or less. In order to proceed utilizing beta regression, we removed these 171 points from our dataset. Another way to address the zero rates without losing any data point is to transform the response

variable y into open interval (0, 1) by taking

$$y' = \frac{y(N-1)+0.5}{N} \tag{4.9}$$

where N is the sample size, proposed by Smithson and Verkuilen [72]. The rationale for this rescaling is presented in their paper as well. The regression models using transformed response variable perform worse than the models without zero rates. Therefore, in this chapter we only present the regression results based on the first method, which removes observations with "zero" (or removed) penetration rates.

For MAPD plans, the beta regression is applied to all data points (as opposed to splitting the data in a training and testing data set to create then test our model) since our purpose in using analytics here is descriptive rather than predictive or prescriptive in nature. Logit, probit, cloglog and loglog link functions are used. For MAPD plans, the loglog link function performs the best with highest pseudo R^2 and smallest AIC (Akaike Information Criterion). Therefore, we choose the loglog link function to proceed with the model build.

Table 4.2: Comparisons of results using different link functions

Link function	Log-likelihood	Pseudo R^2	AIC
Logit	3510	0.4837	-7002.891
Probit	3528	0.5081	-7032.214
Cloglog	3486	0.4712	-6957.166
Loglog	3557	0.5257	-7076.105

We first include all potential independent variables in the beta regression model; however, some of the variables such as the percentage of people with heart attack are not statistically significant in the full model. We thus delete all non-significant variables and fit the model again until all variables left in the model are significant at the 0.05 significance level. The regression coefficients are illustrated in table 4.3.

The significance codes are: 0.000001 '***', 0.001 '**', 0.01 '*', 0.05 '.', 0.1 ' '.

Table 4.3: Parameter estimates - MAPD

Parameter	Estimate	Std. error	z value	Pr(> |z|)	Sig. code
(Intercept)	2.4790	0.3156	7.855	0.0000	***
cnt_mapd	0.0208	0.0007	29.439	0.0000	***
Average_Age	-0.0508	0.0042	-12.04	0.0000	***
Percent_Female	-0.8483	0.2964	-2.862	0.0042	**
Percent_Non_Hispanic_White	0.1874	0.0388	4.831	0.0000	***
Bachelor_degree_or_higher_perc	-0.4411	0.0790	-5.581	0.0000	***
Homeownership_rate	0.4534	0.0778	5.83	0.0000	***
Percent_Eligible_for_Medicaid	-0.2359	0.0961	-2.454	0.0141	*
Average_HCC_Score	1.5523	0.1310	11.85	0.0000	***
Percent_of_kidney_disease	0.9764	0.2592	3.767	0.0002	***
Percent_of_obstructive_pulmonary	-2.2263	0.2395	-9.295	0.0000	***
Percent_of_diabetes	-1.3787	0.2214	-6.226	0.0000	***
Percent_of_ischemic_heart	-1.4635	0.1359	-10.768	0.0000	***
Percent_of_breast_cancer	-1.7833	0.9687	-1.841	0.0656	.
Percent_of_colorectal_cancer	-3.6122	1.2072	-2.992	0.0028	**
Percent_of_hypertension	-1.1590	0.1552	-7.467	0.0000	***
Percent_of_high_cholesterol	0.5189	0.1105	4.696	0.0000	***
Percent_of_alzheimer	0.9932	0.3452	2.877	0.0040	**
Part_B_Drugs_Standardized_Costs_perc	0.8453	0.3578	2.362	0.0182	*
Percent_of_Beneficiaries_Using_PB	0.2277	0.0557	4.086	0.0000	***
ϕ	26.8666	0.7452	36.05	0.0000	***

Several diagnostic measures are provided to measure the goodness-of-fit of the model and influential observations. The log-likelihood is 3999, and AIC is -7956.1. The pseudo R^2 of the model for MAPD plans is 0.527, where pseudo R^2 is defined as the square of the sample correlation coefficient between $\hat{\eta}$ and $g(y)$ and is thus very similar to the R^2 in OLS linear regression. All explanatory variables are significant at the 0.05 significance level. The diagnostic plot in figure 4.2 also shows the model fits quite well except for a few outliers. The parameter estimates and pseudo R^2 exhibit little change when we fit the model without the outliers.

The model can be specified as

$$g(\mu_t) = \beta_0 + \sum_{k=1}^{19} \beta_k x_{tk} \tag{4.10}$$

where β_k are the coefficients we get from the beta regression, and x_{tk} are the 19 independent variables we select to include in the model. Coefficients estimates in table 4.3 quantify the relationship between each independent variable and the response variable. For instance, there is a positive relationship between average HCC score and MAPD plan penetration. The average HCC score indicates the average health status of Medicare beneficiaries. A higher average HCC score means worse health status of the county's Medicare population. Take variable "cnt_mapd" for example, which is the number of MAPD plans available in the county. The coefficient estimate of this variable in the beta regression can be interpreted as follows. Suppose that the number of MAPD plans available increases by 1, then the log of the ratio between chances of enrolling in a MAPD plan under the new setting relative to the old setting, all other variables remaining the same, is 0.0208. ϕ is the precision parameter estimated from beta regression which defines the shape the beta distribution of the response variable.

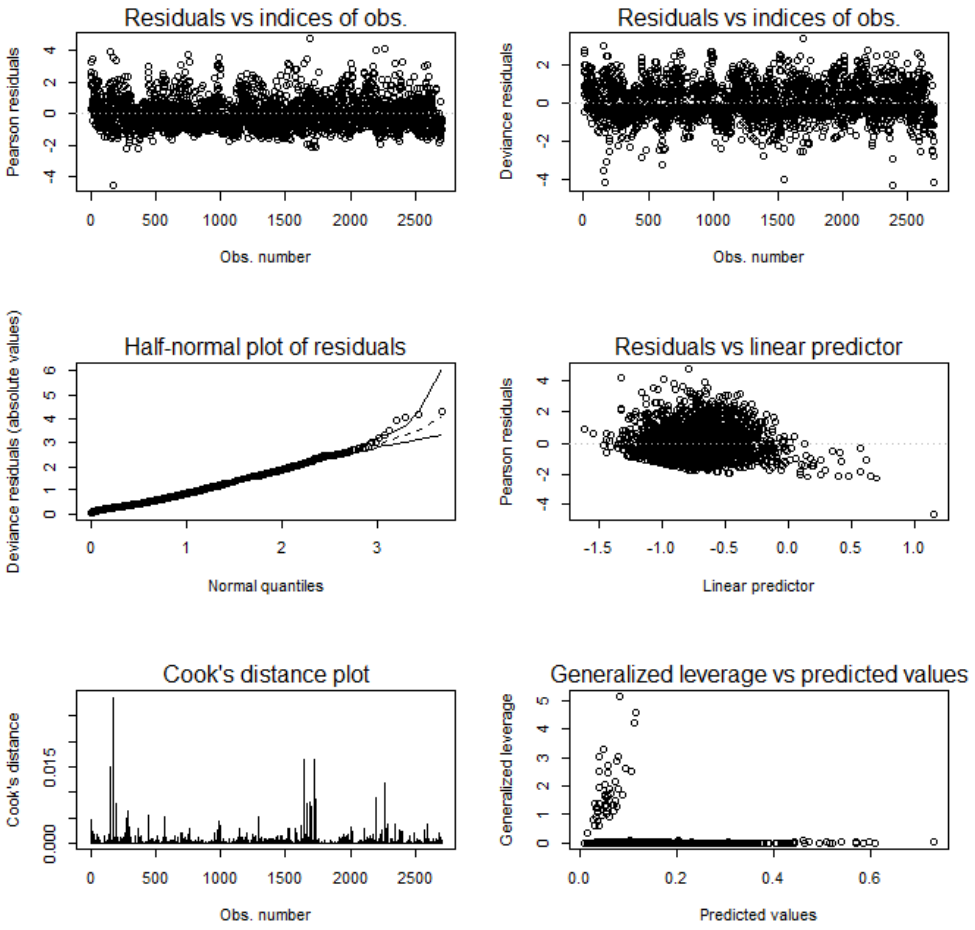


Figure 4.2: Diagnostic plots of MAPD plans

As in the model for MAPD plans, the loglog link function performs best for PDP plan penetration rates and total penetration rates.

Table 4.4: Summary of diagnostic statistics for MAPD, PDP, and Total

Model	Link function	Log-likelihood	Pseudo R^2	AIC
MAPD	**loglog**	**3999**	**0.527**	**-7956.1**
	logit	3533	0.4837	-7002.89
	probit	3548	0.5081	-7032.21
	cloglog	3511	0.4712	-6957.17
PDP	**loglog**	**3210**	**0.5795**	**-6369.2**
	logit	2891	0.5513	-5717.93
	probit	2887	0.5527	-5710.76
	cloglog	2899	0.5439	-5734.13
Total	**loglog**	**3687**	**0.3683**	**-7330.4**
	logit	3334	0.356	-6603.09
	probit	3333	0.356	-6602.03
	cloglog	3330	0.3557	-6576.76

All models' pseudo R^2's are reasonably high, especially the model for the PDP plans, and the diagnostic plots also indicate that the models fit well. To address the potential problem of overfitting, we also split the whole dataset into training set and test set by the 70/30 rule. The model is built on the training set, and the regression result is applied to the test set. The model based on training data has very similar pseudo R^2's for all models. And the MSE's (Mean Squared Error) of the predictions on test data are about 0.01, which is acceptable low to justify that our model is properly fitted. Table 4.5 summarizes the coefficients estimates of significant variables for the three models. Note that N/A means the variable is not significant in the model. Also, a few variables show coefficients as zero. However, the actual coefficient estimate of these variables are very small numbers close to zero, but not zero exactly. Details of regression results are in the appendix.

Table 4.5: Summary of coefficients estimates for MAPD, PDP, and Total

	PDP		MAPD		Total	
Parameter	Estimate	Signi.	Estimate	Signi.	Estimate	Signi.
(Intercept)	-5.9280	***	2.4790	***	-2.9850	***
cnt_mapd	-0.0118	***	0.0208	***	0.0037	***
cnt_pdp	0.0143	***	N/A	N/A	0.0113	***
Average_Age	0.0575	***	-0.0508	***	0.0206	***
Percent_Female	3.1290	***	-0.8483	**	2.3480	***
Percent_Non_Hispanic_White	0.1679	***	0.1874	***	0.3108	***
Bachelor_degree_or_higher_perc	-0.4015	***	-0.4411	***	-0.8792	***
Homeownership_rate	-0.3007	***	0.4534	***	N/A	N/A
Per_capita_money_income	0.0000	***	N/A	N/A	0.0000	.
Persons_below_poverty_level_per	0.8762	***	N/A	N/A	0.5370	***
Percent_Eligible_for_Medicaid	1.5300	***	-0.2359	*	1.5880	***
Average_HCC_Score	-1.1170	***	1.5523	***	0.2831	**
Percent_of_atrial_fibrillation	0.8717	.	N/A	N/A	1.5850	***
Percent_of_depression	N/A	N/A	N/A	N/A	0.3751	*
Percent_of_heart_failure	N/A	N/A	N/A	N/A	-0.3908	*
Percent_of_kidney_disease	-0.6750	**	0.9764	***	N/A	N/A
Percent_of_obstructive_pulmonary	1.1110	***	-2.2263	***	-1.1500	***
Percent_of_diabetes	1.3410	***	-1.3787	***	N/A	N/A
Percent_of_ischemic_heart	0.8841	***	-1.4635	***	N/A	N/A
Percent_of_breast_cancer	N/A	N/A	-1.7833	.	N/A	N/A
Percent_of_colorectal_cancer	4.6670	***	-3.6122	**	3.8520	***
Percent_of_prostate_cancer	1.7760	*	N/A	N/A	N/A	N/A
Percent_of_asthma	-1.7180	***	N/A	N/A	-2.2520	***
Percent_of_hypertension	1.0210	***	-1.1590	***	N/A	N/A
Percent_of_high_cholesterol	-0.8609	***	0.5189	***	-0.4981	***
Percent_of_alzheimer	-1.3240	***	0.9932	**	N/A	N/A
Percent_of_stroke	-2.6300	***	N/A	N/A	-3.2040	***
Part_B_Drugs_Standardized_Costs_perc	N/A	N/A	0.8453	*	0.5984	.
Percent_of_Beneficiaries_Using_PB	N/A	N/A	0.2277	***	0.1324	**

All models share most variables, with some exceptions where some variables are significant in one model but might not be significant in others. For example, the variable that indicates percentage of people who have asthma has negative impact on PDP penetration but has no significant impact on MAPD penetration. This means that counties where fewer people have asthma tend to enroll in PDP plans. In contrast, the variable about the percentage of beneficiaries using Part B has a positive relationship with MAPD plan penetration, but not PDP plan penetration. This makes sense since it is required to enroll in both Part A and Part B in order to

enroll in MAPD plans. Similar factors include number of PDP plans available, per capita money income, percentage of people below poverty level, percentage of people with atrial fibrillation/breast cancer/prostate cancer/stroke, and Part B drug standardized Medicare costs as a percentage of total standardized Medicare costs. Two variables have the same positive/negative signs of their coefficient estimates: percentage of non-Hispanic white population, and percentage of bachelor degree or higher population. This suggests that these two factors are affecting MAPD enrollments and PDP enrollments in the same direction. Fourteen variables have opposite signs in the coefficients for MAPD and PDP plan penetration. Nine factors that have positive effects on PDP penetration rates but negative effects on MAPD penetration rates are: number of MAPD plans available, average age, percentage of female, percentage of people eligible for Medicaid, and percentage of people with obstructive pulmonary/ diabetes/ ischemic heart/ colorectal cancer/ hypertension. While five factors that are negatively related to PDP enrollments but positively related to MAPD enrollments include homeownership rate, average HCC score, percentage of people with kidney disease/high cholesterol/ Alzheimer.

4.6 Conclusions

This chapter identifies factors associated with the choice of the Medicare population between MAPD plans and PDP plans when they receive prescription drug coverage. The data in the analysis includes high-level census data, plan data and penetration data. We use beta regression due to the properties of the response variable and implement the methodology in R. We find that half of the variables have opposite signs, which means they have impacts on MAPD and PDP plan penetration rates in different directions. This conclusion can help policy-makers identify which types of

prescription drug plans for the Medicare population should be emphasized in each county based on the characteristics of their population from census data and plan data of the counties.

4.7 Appendix

This appendix includes regression coefficients estimates as well as diagnostic plots for PDP and total penetration models.

Table 4.6: Parameter estimates - PDP

Parameter	Estimate	Std. error	z value	Pr($>\|z\|$)	Sig. code
(Intercept)	-6.1010	0.2965	-20.5770	0.0000	***
cnt_mapd	-0.0120	0.0006	-21.3100	0.0000	***
cnt_pdp	0.0146	0.0024	6.1610	0.0000	***
Average_Age	0.0599	0.0039	15.3470	0.0000	***
Percent_Female	3.1450	0.2649	11.8710	0.0000	***
Percent_Non_Hispanic_White	0.1928	0.0357	5.3950	0.0000	***
Bachelor_degree_or_higher_perc	-0.4200	0.0986	-4.2610	0.0000	***
Homeownership_rate	-0.3101	0.0755	-4.1080	0.0000	***
Per_capita_money_income	0.0000	0.0000	3.5540	0.0004	***
Persons_below_poverty_level_per	0.8545	0.1420	6.0180	0.0000	***
Percent_Eligible_for_Medicaid	1.5290	0.0916	16.6960	0.0000	***
Average_HCC_Score	-1.0820	0.1203	-8.9920	0.0000	***
Percent_of_kidney_disease	-0.6899	0.2363	-2.9200	0.0035	**
Percent_of_obstructive_pulmonary	1.1040	0.2215	4.9840	0.0000	***
Percent_of_diabetes	1.2640	0.2011	6.2880	0.0000	***
Percent_of_ischemic_heart	0.8847	0.1246	7.1000	0.0000	***
Percent_of_colorectal_cancer	5.0150	1.0940	4.5830	0.0000	***
Percent_of_prostate_cancer	2.1120	0.6920	3.0530	0.0023	**
Percent_of_asthma	-1.8460	0.4575	-4.0350	0.0001	***
Percent_of_hypertension	1.0150	0.1466	6.9250	0.0000	***
Percent_of_high_cholesterol	-0.8301	0.1044	-7.9500	0.0000	***
Percent_of_alzheimer	-1.2590	0.3161	-3.9830	0.0001	***
Percent_of_stroke	-2.4560	0.6370	-3.8560	0.0001	***
ϕ	45.362	1.223	37.09	0.0000	***

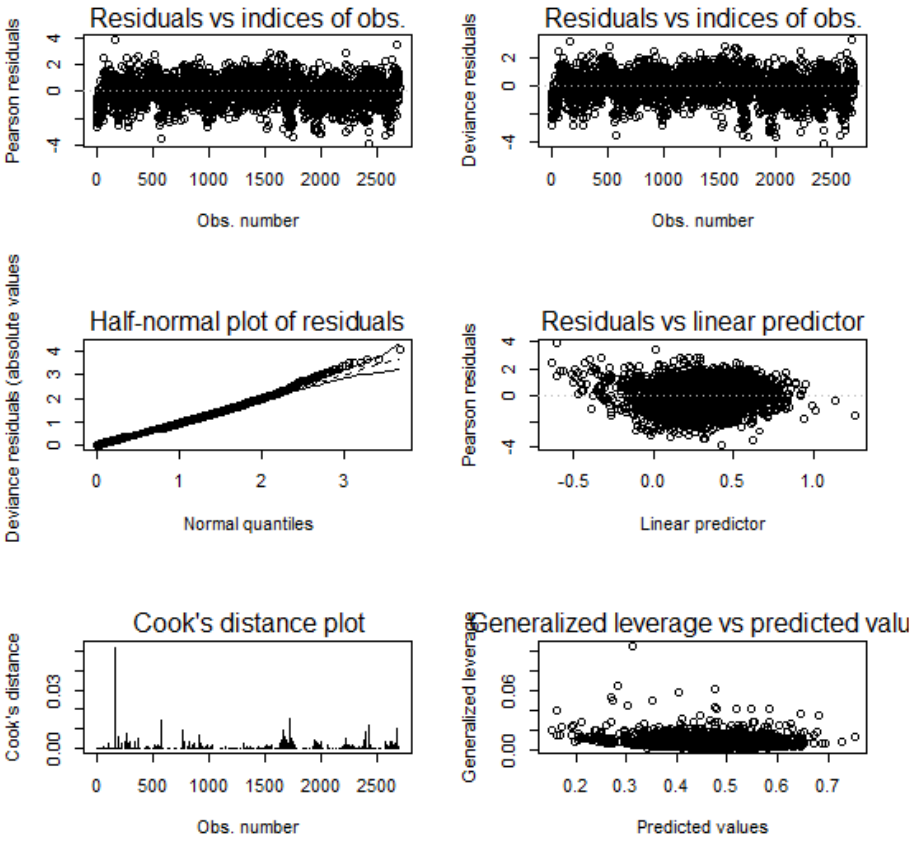


Figure 4.3: Diagnostic plots of PDP plans

Table 4.7: Parameter estimates - Total

Parameter	Estimate	Std. error	z value	Pr($>\|z\|$)	Sig. code
(Intercept)	-2.7131	0.2611	-10.3920	0.0000	***
cnt_mapd	0.0039	0.0006	6.9630	0.0000	***
cnt_pdp	0.0107	0.0023	4.7130	0.0000	***
Average_Age	0.0178	0.0036	4.9080	0.0000	***
Percent_Female	2.3613	0.2222	10.6260	0.0000	***
Percent_Non_Hispanic_White	0.3362	0.0341	9.8450	0.0000	***
Bachelor_degree_or_higher_perc	-0.7069	0.0567	-12.4630	0.0000	***
Persons_below_poverty_level_per	0.3744	0.1006	3.7220	0.0002	***
Percent_Eligible_for_Medicaid	1.5960	0.0878	18.1850	0.0000	***
Average_HCC_Score	0.2620	0.0882	2.9720	0.0030	**
Percent_of_atrial_fibrillation	1.7105	0.4143	4.1290	0.0000	***
Percent_of_obstructive_pulmonary	-1.2313	0.1863	-6.6100	0.0000	***
Percent_of_colorectal_cancer	3.6461	1.0294	3.5420	0.0004	***
Percent_of_asthma	-2.2510	0.4292	-5.2450	0.0000	***
Percent_of_high_cholesterol	-0.4216	0.0685	-6.1510	0.0000	***
Percent_of_stroke	-3.4238	0.5715	-5.9910	0.0000	***
Percent_of_Beneficiaries_Using_PB	0.1279	0.0459	2.7880	0.0053	**
ϕ	63.67	1.72	37.01	0.0000	***

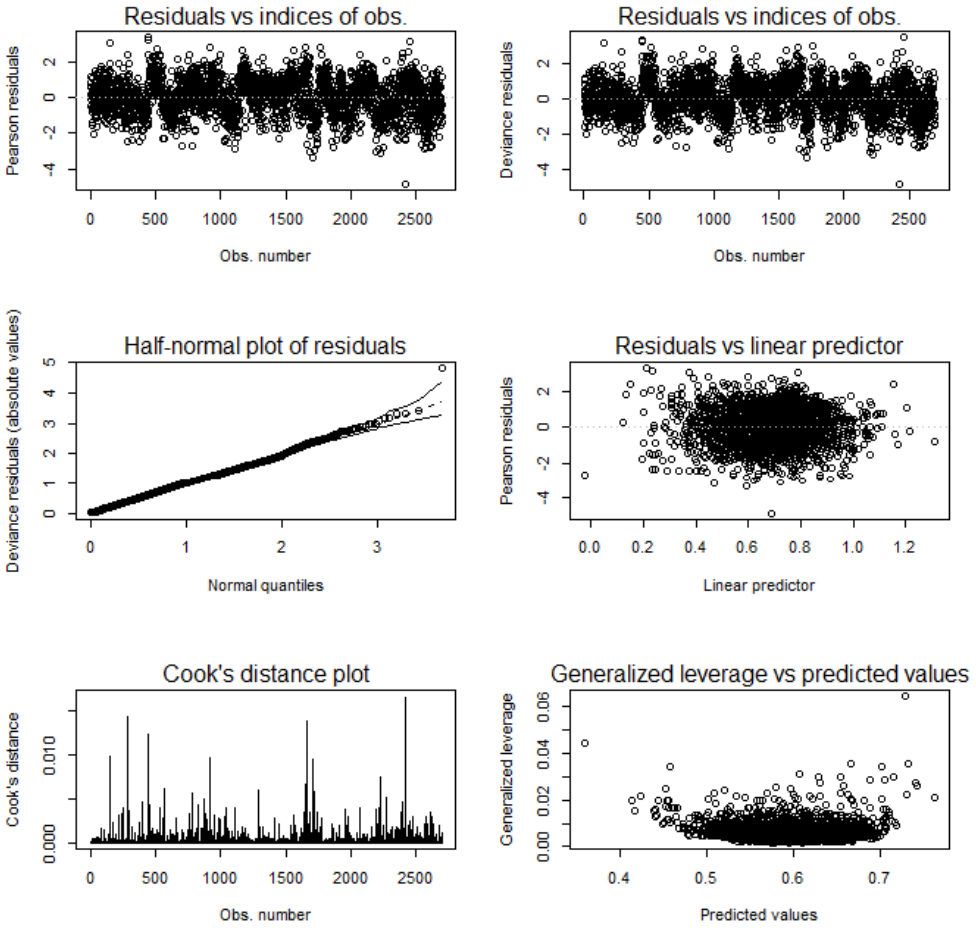


Figure 4.4: Diagnostic plots of all Part D plans

Chapter 5

Analysis of Medicare Services Usage and Reimbursement Rate

5.1 Introduction

The Medicare program pays healthcare costs for the elderly, the permanently disabled, and those with end-stage renal disease. In 2013, Medicare spending accounted for 14% of the federal budget as shown in Figure 5.1. It also accounted for 20% of total national health spending in 2012, 27% of spending on hospital care, and 23% of spending on physician services. Medicare benefit payments totaled $583 billion in 2013. About 1/4 was for hospital inpatient services, 12% for physician services, and 11% for the Part D drug benefit. Another 1/4 of benefit spending was for Medicare Advantage private health plans covering all Part A and B benefits. In 2014, 30% of Medicare beneficiaries are enrolled in Medicare Advantage plans [44].

CMS (Center for Medicare and Medicaid Services) has recently made available

extensive data files about providers' utilization profiles, i.e., what they are charging Medicare for, for each provider, all over the country, for services provided at a facility or at an office. The objective of this chapter is to get statistical insights about the health of the Medicare population from the public files. We are interested in answering the following questions particularly: Are there states with a disproportionate incidence of a given HCPCS (Healthcare Common Procedure Coding System) code compared to the size of Medicare population in that state? Are usage disparities across states driven by differences in Medicare reimbursement amounts?

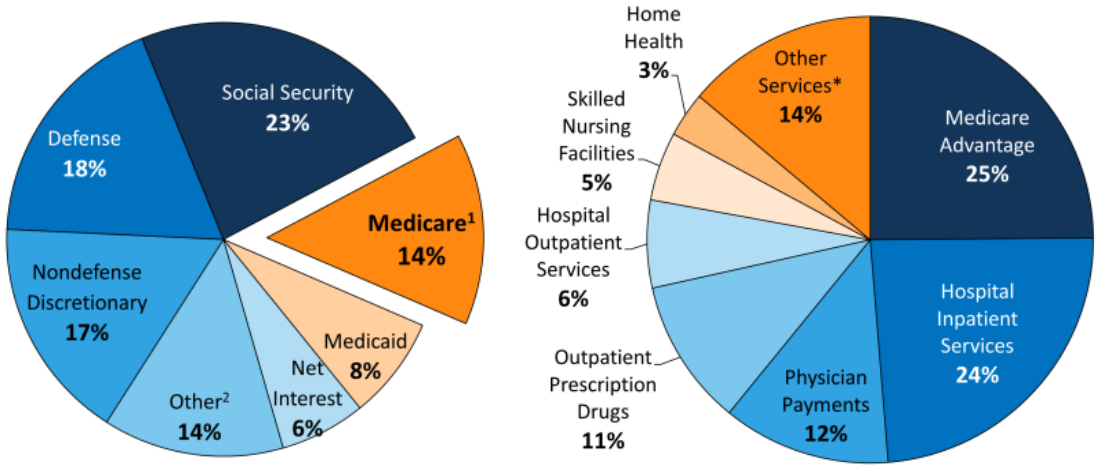


Figure 5.1: Medicare vs. Federal Budget & Medicare Benefit Payments, 2013

Source: Kaiser Family Foundation [44]

5.2 Data Description

As part of the Obama Administrations efforts to make the healthcare system more transparent, affordable, and accountable, the Centers for Medicare & Medicaid Services has prepared a public data set, the Medicare Provider Utilization and Payment Data [41], with information on services and procedures provided to Medicare beneficiaries by physicians and other healthcare professionals. The summary data contains

aggregated information by state and HCPCS code. Any aggregated records which are derived from 10 or fewer beneficiaries are excluded to protect the privacy of Medicare beneficiaries. A total of 8305 HCPCS codes and 880644 physicians in 50 states, District of Columbia, Puerto Rico and Virgin Islands are included in the data. The data includes physician services (Part B) only. Medicare Part B beneficiaries represent about 90 percent of total Medicare population (Part A and Part B). Table 5.1 illustrates the main information provided in the data. We will analyze the data to get insights on physician services usage and Medicare reimbursement rate. All data analyses in this chapter are performed in SAS.

Table 5.1: Medicare Provider Utilization and Payment Data, 2012

Variable Name	Description	Example
nppes_provider_state_desc	The state where the provider is located.	Pennsylvania
hcpcs_code	HCPCS code for the service provided.	67028
hcpcs_description	Description of the HCPCS code.	Injection eye drug
place_of_service	Either a facility ('F') or non-facility ('O').	O
number_of_providers	Number of providers within state.	147
line_srvc_cnt	Number of services provided.	103114
unique_bene_doctor_cnt	Number of unique beneficiary/doctor interactions.	25393
bene_day_srvc_cnt	Number of distinct Medicare beneficiary/per day services.	102732
average_medicare_allowed_amt	Average of the Medicare allowed amount for the service.	$119.847719
average_submitted_chrg_amt	Average of the charges submitted for the service.	$604.2012631
average_medicare_payment_amt	Average amount that Medicare paid.	$93.4530861

5.3 Correlation Analysis of Services Usage and Medicare Population/ Reimbursement

We first create an additional column - total dollar amount - by multiplying the Medicare average allowed amount by the line service count. Then we rank the data in decreasing order by this column. The top ten HCPCS codes in total dollar amount are shown in table 2. We notice that the most expensive codes in total dollar amount are 99214 (office/outpatient visit, medium level of complexity, in an office setting),

99213 (office/outpatient visit, low level of complexity, in an office setting), and 99232 (subsequent hospital care, in facility).

Table 5.2: Top 10 HCPCS Codes in Total Dollar Allowed Amount

	State	HCPCS	HCPCS Desc	Place	# of Srvc	Avg Allowed Amt	Tot Allowed Amt
1	Overall	99214	Office/outpatient visit est	O	76,237,819	$103.11	$7,861,014,700
2	Overall	99213	Office/outpatient visit est	O	91,252,355	$69.59	$6,350,698,955
3	Overall	99232	Subsequent hospital care	F	49,384,357	$69.87	$3,450,330,601
4	Overall	66984	Cataract surg w/iol 1 stage	F	3,716,637	$611.76	$2,273,702,288
5	Overall	99223	Initial hospital care	F	11,213,325	$195.75	$2,195,032,129
6	Overall	99233	Subsequent hospital care	F	21,423,718	$100.73	$2,157,916,561
7	Overall	A0427	ALS1-emergency	F	4,980,848	$413.56	$2,059,889,581
8	Overall	99285	Emergency dept visit	F	10,317,768	$167.62	$1,729,505,516
9	Overall	A0428	bls	F	6,587,015	$221.67	$1,460,121,921
10	Overall	92014	Eye exam & treatment	O	11,217,016	$118.05	$1,324,162,376

We conduct correlation analyses on these three codes to investigate if the incidence of codes per state more strongly correlated with Medicare population, or with Medicare reimbursement in that state. Table 5.3 shows the Pearson correlation coefficients for code 99214. Service usage and Medicare population are highly correlated as their correlation is 0.96, with p-value <0.0001. Meanwhile, the correlation between service usage and Medicare reimbursement is only 0.21, which indicates that services usage is less correlated with reimbursement rate than population. Same conclusion can be drawn from the scatter plots as it is close to a straight line in service usage versus population plot but disperse in service usage versus reimbursement plot.

Table 5.3: Pearson Correlation Coefficients for 99214

Variable	With Variable	N	Correlation	95% Confidence Interval	p Value
line_srvc_cnt	enrl_partb	53	0.96305	[0.935392,0.978209]	<.0001
line_srvc_cnt	avg_allowed_amt	53	0.21008	[-0.065851,0.452958]	0.1316

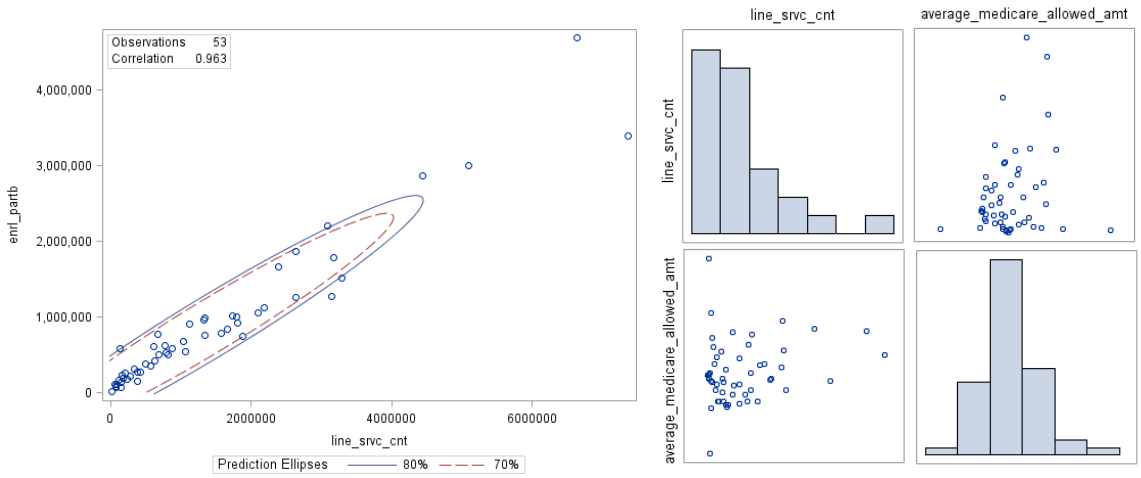


Figure 5.2: Scatter Plots of 99214

We can get similar conclusions for code 99213 and code 99232. Their Pearson correlation coefficients are shown in table 5.4 and table 5.5, respectively.

Table 5.4: Pearson Correlation Coefficients for 99213

Variable	With Variable	N	Correlation	95% Confidence Interval	p Value
line_srvc_cnt	enrl_partb	53	0.97857	[0.962296,0.987399]	<.0001
line_srvc_cnt	avg_allowed_amt	53	0.22889	[-0.04633,0.468384]	0.0994

Table 5.5: Pearson Correlation Coefficients for 99232

Variable	With Variable	N	Correlation	95% Confidence Interval	p Value
line_srvc_cnt	enrl_partb	53	0.94172	[0.898962,0.965492]	<.0001
line_srvc_cnt	avg_allowed_amt	53	0.16224	[-0.11455,0.413069]	0.2471

5.4 Most Expensive/Common Services in Each State

In order to make recommendations to the policy makers which line items to pick in each state to try to decrease the dollar amount associated by reducing incidence

numbers (not reimbursement levels), we first identify the most expensive and most common line items in each state. We take Pennsylvania as an example. The Medicare population (part B) in Pennsylvania is 2,200,614 in 2012. Per_capita is calculated as total dollar amount divided by Medicare population in PA. The three most expensive codes in PA are the same as in the United State as a whole, while the rest vary from their country level counterparts.

Table 5.6: Ten most expensive HCPCS codes in PA

hcpcs_code	hcpcs_description	place of srvc	total_dollar	per_capita
99214	Office/outpatient visit est	O	$ 316,631,372	$ 143.88
99213	Office/outpatient visit est	O	$ 239,046,599	$ 108.63
99232	Subsequent hospital care	F	$ 188,314,151	$ 85.57
A0428	BLS	F	$ 97,205,563	$ 44.17
99223	Initial hospital care	F	$ 95,920,544	$ 43.59
J2778	Ranibizumab injection	O	$ 89,023,943	$ 40.45
99285	Emergency dept visit	F	$ 88,333,927	$ 40.14
66984	Cataract surg w/iol 1 stage	F	$ 85,919,141	$ 39.04
A0427	ALS1-emergency	F	$ 81,925,077	$ 37.23
99233	Subsequent hospital care	F	$ 69,563,073	$ 31.61

We do the same calculation for all states and count the frequency a HCPCS code appearing in the top ten most expensive list of each state. The frequency counts are shown in table 5.7. For example, code 99213 and 99214 appear in every single state's (including Puerto Rico and Virgin Islands) top ten most expensive list. However, for some codes such as 83898 (Molecule nucleic ampli each) only happens to Utah, not any other state.

Table 5.7: Most expensive HCPCS codes over all states

hcpcs code	hcpcs desc	place of service	cnt_state
99213	Office/outpatient visit est	O	53
99214	Office/outpatient visit est	O	53
99232	Subsequent hospital care	F	51
66984	Cataract surg w/iol 1 stage	F	49
A0427	ALS1-emergency	F	48
99223	Initial hospital care	F	43
99285	Emergency dept visit	F	40
99233	Subsequent hospital care	F	38
92014	Eye exam & treatment	O	28
97110	Therapeutic exercises	O	18
A0425	Ground mileage	F	16
A0428	BLS	F	15
J2778	Ranibizumab injection	O	13
A0429	BLS-emergency	F	10

Table 5.8: Unique most expensive HCPCS codes

hcpcs code	hcpcs desc	place of service	state_name
83898	Molecule nucleic ampli each	O	Utah
83904	Molecule mutation identify	O	Utah
83909	Nucleic acid high resolute	O	Utah
83914	Mutation ident ola/sbce/aspe	O	Washington
90960	Esrd srv 4 visits p mo 20+	O	Puerto Rico
97140	Manual therapy	O	Wyoming
99212	Office/outpatient visit est	O	Wyoming
99222	Initial hospital care	F	Michigan
A0430	Fixed wing air transport	F	Alaska
A0431	Rotary wing air transport	F	Alabama
A0435	Fixed wing air mileage	F	Alaska
G9152	Mapcp demo community	O	Vermont

For the most common HCPCS codes identification, we still take Pennsylvania as an example, as shown in table 5.9. We rank the data in decreasing order by service usage. We notice that despite being highest in total dollar amount, code 99214 is not the highest in terms of service usage.

Table 5.9: Ten most common HCPCS codes in PA

hcpcs_code	hcpcs_description	place of service	line_srvc_cnt
P9603	One-way allow prorated miles	O	6,465,047
A0425	Ground mileage	F	5,989,114
J0881	Darbepoetin alfa, non-esrd	O	3,585,133
99213	Office/outpatient visit est	O	3,453,943
99214	Office/outpatient visit est	O	3,084,848
99232	Subsequent hospital care	F	2,711,119
36415	Routine venipuncture	O	2,388,292
97110	Therapeutic exercises	O	1,867,715
J1756	Iron sucrose injection	O	1,476,971
Q9967	LOCM 300-399mg/ml iodine,1ml	O	1,413,515

We identify the most common HCPCS codes over all states in the same fashion as identifying most expensive ones. We notice that in table 5.10, code 99232 drops from the third place on the "most expensive" list to the sixth place here, which indicates that this code is significantly expensive.

Table 5.10: Most common HCPCS codes over all states

hcpcs_code	hcpcs_description	place of service	cnt_state
99213	Office/outpatient visit est	O	53
99214	Office/outpatient visit est	O	53
A0425	Ground mileage	F	53
Q9967	LOCM 300-399mg/ml iodine,1ml	O	48
36415	Routine venipuncture	O	46
99232	Subsequent hospital care	F	40
97110	Therapeutic exercises	O	34
J0881	Darbepoetin alfa, non-esrd	O	28
P9603	One-way allow prorated miles	O	14
85025	Complete cbc w/auto diff wbc	O	13
J0878	Daptomycin injection	O	13
Q0138	Ferumoxytol, non-esrd	O	12
80053	Comprehen metabolic panel	O	11
97140	Manual therapy	O	11

In the unique most common list (table 5.11), code 83898 (Molecule nucleic ampli

each) appears again for Utah, which might be an indicator that Utah should focus more on this code than other states.

Table 5.11: Unique most common HCPCS codes

hcpcs code	hcpcs desc	place of service	state name
66984	Cataract surg w/iol 1 stage	O	Kentucky
81000	Urinalysis nonauto w/scope	O	Virgin Islands
83896	Molecular diagnostics	O	Virginia
83898	Molecule nucleic ampli each	O	Utah
83901	Molecule nucleic ampli addon	O	Washington
83904	Molecule mutation identify	O	Utah
83909	Nucleic acid high resolute	O	Utah
83914	Mutation ident ola/sbce/aspe	O	Washington
83925	Assay of opiates	O	Rhode Island
84443	Assay thyroid stim hormone	O	Puerto Rico
98940	Chiropractic manipulation	O	North Dakota
99212	Office/outpatient visit est	O	Alaska
99231	Subsequent hospital care	F	Louisiana
G0008	Admin influenza virus vac	O	Arizona
G9151	Mapcp demo state	O	Michigan
G9153	Mapcp demo physician	O	Michigan
J1170	Hydromorphone injection	O	Mississippi
J1745	Infliximab injection	O	Montana
J2323	Natalizumab injection	O	Utah
J3010	Fentanyl citrate injeciton	O	Mississippi
L8621	Repl zinc air battery	O	Colorado

Combining information about per capita and number of services, we would recommend targeting specific line times to decrease total dollar amount by reducing incidence numbers. For instance, the recommendation for Pennsylvania would be focusing first on codes 99214, 99213 and 99232 since they have the highest total dollar amounts and relatively intense service usage.

Table 5.12: Recommendation for Pennsylvania

hcpcs	hcpcs_description	place	per_capita	line_srvc_cnt	total_dollar
P9603	One-way allow prorated miles	O	$ 0.90	6,465,047	$ 1,981,847
A0425	Ground mileage	F	$ 20.63	5,989,114	$ 45,392,499
J0881	Darbepoetin alfa, non-esrd	O	$ 5.33	3,585,133	$ 11,725,185
99213	Office/outpatient visit est	O	$ 108.63	3,453,943	$ 239,046,599
99214	Office/outpatient visit est	O	$ 143.88	3,084,848	$ 316,631,372
99232	Subsequent hospital care	F	$ 85.57	2,711,119	$ 188,314,151
36415	Routine venipuncture	O	$ 3.26	2,388,292	$ 7,164,785
97110	Therapeutic exercises	O	$ 23.98	1,867,715	$ 52,762,284
J1756	Iron sucrose injection	O	$ 0.20	1,476,971	$ 435,796
Q9967	LOCM 300-399mg/ml iodine,1ml	O	$ 0.09	1,413,515	$ 193,339

5.5 Services and States Most Related to Medicare Reimbursement

There is concern that doctors and hospitals tend to upcode on purpose to get more reimbursement from Medicare [70]. To investigate whether this concern is valid or not, we analyze the correlation between service codes and Medicare reimbursement over all states. For service codes, we analyze the top 1000 codes in total dollar amount as it is cost efficient. Service usage is surprisingly not highly correlated with reimbursement. Among 1000 codes, only 9 of them have correlation higher than 0.7 and 7 of them higher than 0.5 but lower than 0.7. Some of the most highly correlated codes are shown in table 5.13. The concern is valid since it does exist that some expensive codes are more often used than others. Medicare could pay more attention to these codes to prevent upcoding.

Table 5.13: Services most related to Medicare reimbursement

hcpcs_code	hcpcs_description	place of service	correlation
89240	Pathology lab procedure	O	1
0182T	Hdr elect brachytherapy	O	0.983123
84999	Clinical chemistry test	O	0.870835
G0249	Provide INR test mater/equip	O	0.836478
86849	Immunology procedure	O	0.741178

We do a similar analysis on the codes at state level. The results show that service usage doesn't seem to be correlated to the Medicare allowed reimbursement on state level as even the most significant correlation between number of services and allowed amount is only -0.08. This means that no state particularly upcodes.

5.6 Simulation of Project Growth in Usage and Spending

Our objective in this section is to project growth in service usage and spending in ten years. We get the total number of Medicare beneficiaries by state from CMS's Medicare Enrollment Reports [40] and Kaiser Family Foundation's State Health Facts [43] due to lack of partial data from CMS's enrollment database. Then we perform a simulation about the projected growth in line item counts (service usage) based on growth in Medicare population for each state in ten years. We predict Medicare spending in ten years as well assuming Medicare reimbursement rates grow at inflation rate. The steps of the prediction process is as follows:

1. Perform time series analysis on Medicare population of each state from 1999 to 2012, and project Medicare population in ten years based on the growth rates

in each state. We choose ARIMA models because they are, in theory, the most general class of models for forecasting a time series.

2. Analyze relationship between service usage and Medicare population using linear regression, and predict growth in service usage based on the linear model. We could also use nonlinear regression, such as Support Vector Machine regression or Neural Network regression.

3. Predict Medicare spending in ten years based on projected growth in service usages and the assumption that reimbursement rates increasing at inflation rate.

We plot the Medicare population for all states from 1985 to 2012. Almost all states' Medicare population grow steadily over years, except the District of Columbia. The figures below show the Medicare population growth trend for Pennsylvania and U.S. as a whole.

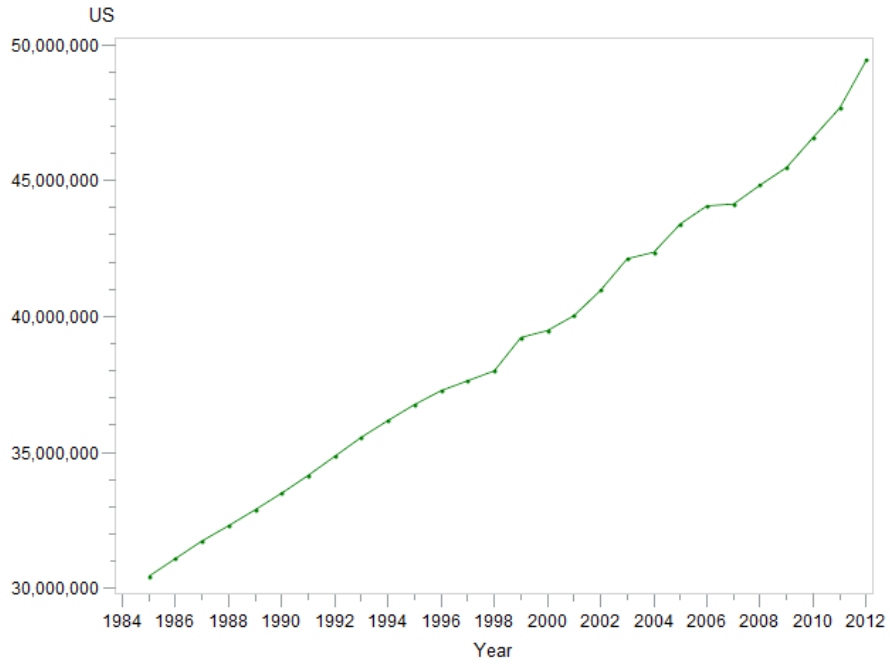


Figure 5.3: US Medicare Population, 1985-2012

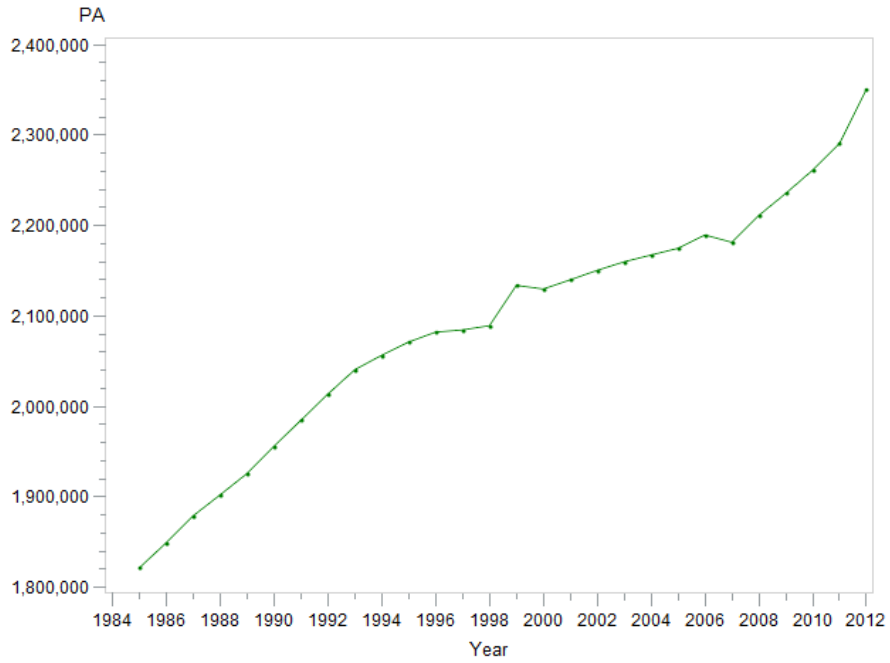


Figure 5.4: PA Medicare Population, 1985-2012

We take PA as an example showing how time series analysis works on population growth projection. In order to decide which time series model to apply on PA population data, we first check autocorrelation of the series.

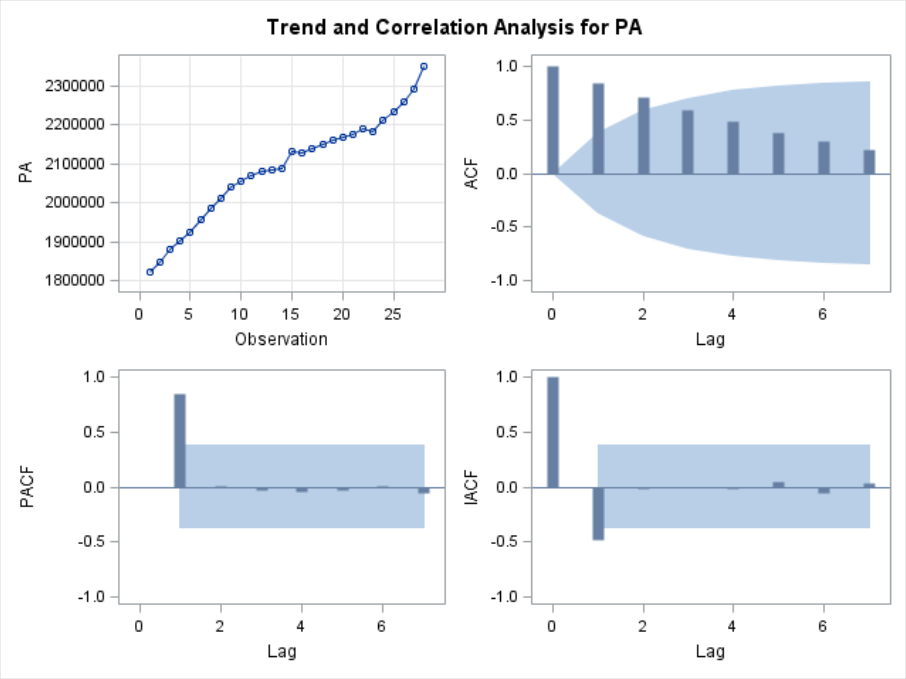


Figure 5.5: Autocorrelation check PA

This series is obviously nonstationary as it has a strong upward trend. Its ACF decays slowly and its PACF only has one spike at lag 1. This is an indicator that AR(1) model might be adequate for the data. Thus, we difference the data once, and the total number of observations drops from 28 to 27 since one data is lost in the differencing process.

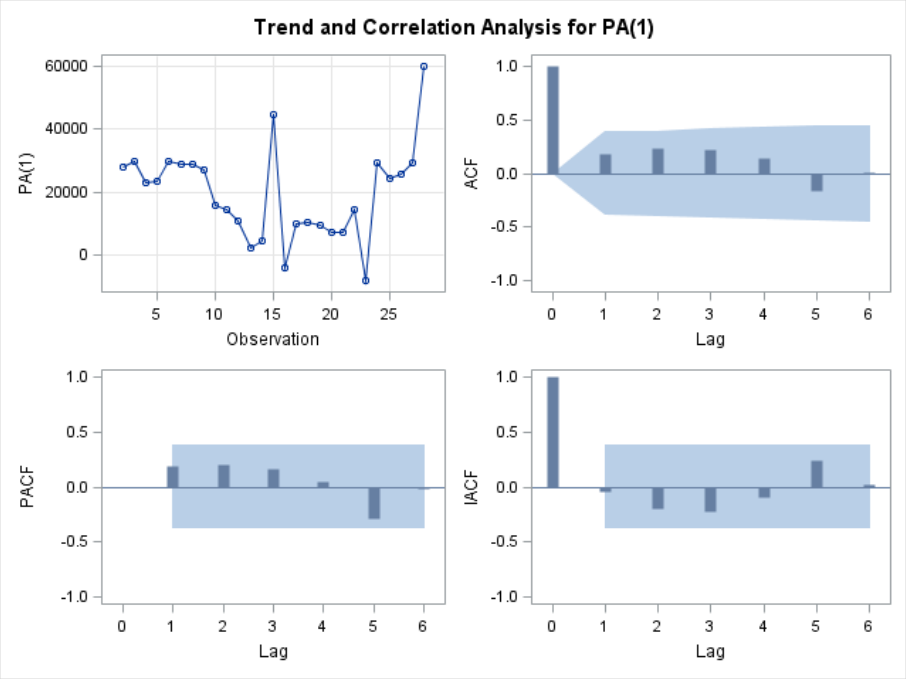


Figure 5.6: Autocorrelation check PA(1)

The time series looks stationary now after differenced once. Thus, AR(1) is adequate for our PA time series data. We apply AR(1) model on the data and forecast the Medicare population in PA in ten years. The forecasts are shown in table 5.14.

Table 5.14: Forecasts for Medicare population in PA

Year	Forecast	Std Error	95% Confidence Limits
2013	2381386	14596.31	[2352778, 2409994]
2014	2404455	23542.82	[2358312, 2450598]
2015	2425464	30570.45	[2365547, 2485381]
2016	2445926	36408.4	[2374567 , 2517285]
2017	2466243	41466.43	[2384970, 2547515]
2018	2486521	45979.6	[2396402, 2576639]
2019	2506789	50089.78	[2408614, 2604963]
2020	2527054	53887.86	[2421436 , 2632672]
2021	2547318	57435.46	[2434747 , 2659890]
2022	2567583	60776.36	[2448463, 2686702]

We compare the actual Medicare population (blue) and predicted Medicare population (red). The model seems to fit well. According to the forecast, the Medicare population in PA in ten years would be 2,567,583.

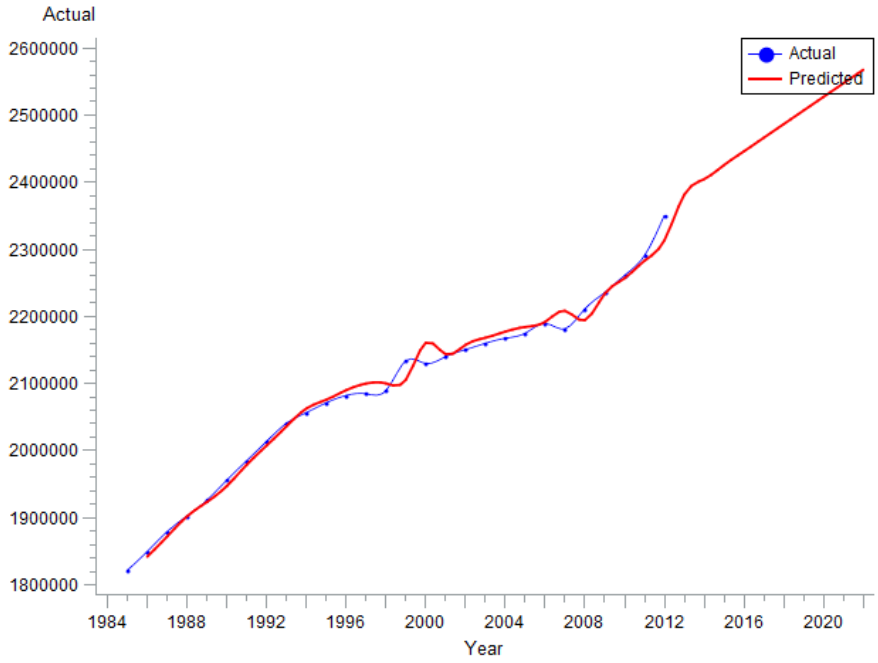


Figure 5.7: Actual vs. Predicted

We apply linear regression on service usage of code 99214 and Medicare population in PA. The R-squared ratio is 0.9275, which is very good. Thus, the projected service usage of code 99214 would be 4,272,833 in ten years. The current service usage of code 99214 is 3,084,848. We can then get the projected spending on this code by multiplying the projected service usage by reimbursement rate, assuming reimbursement rate increases at inflation rate. The current average medicare reimbursement allowed amount for code 99214 in PA is \$102.64. If the inflation rate is 2%, then the Medicare spending on this code in PA would be \$534,606,556 in ten years. We can apply the same methods and get projections for all codes and all states.

5.7 Conclusions

This chapter investigates trends in physician services usage and Medicare reimbursement rate from CMS public files. We analyze the HCPCS (Healthcare Common Procedure Coding System) codes and investigate the validity of the concern that doctors tend to upcode on purpose for more reimbursement. Our findings indicate that services usage is more correlated with population than reimbursement rate, and no state particularly upcodes. In addition, we make recommendations to policy makers on which HCPCS codes to choose for the purpose of reducing Medicare spending. For example, the recommendation for Pennsylvania would be focusing first on codes 99214 (office/outpatient visit, medium level of complexity, in an office setting), 99213 (office/outpatient visit, low level of complexity, in an office setting), and 99232 (subsequent hospital care, in facility) since they have the highest total dollar amounts and relatively intense service usage. We also utilize time series analysis to predict Medicare spending in ten years.

Chapter 6

Robust optimization in Healthcare Systems Engineering - A survey

6.1 Overview of Healthcare Systems in the United States

6.1.1 The Payment System

According to CMS's data [6], U.S. health care spending increased 3.7 percent to reach $2.8 trillion in 2012, an average of $8,915 per person, the fourth consecutive year of slow growth. The share of the economy devoted to health spending decreased from 17.3 percent in 2011 to 17.2 percent in 2012, as the Gross Domestic Product increased nearly one percentage point faster than health care spending at 4.6 percent. Figure 6.1 shows that from 1960 to 2012, the National Health Expenditures and Per Capita Amount have been consistently growing at unsustainable rates over time.

The percentage of GDP devoted to health spending increased rapidly from 1960 to 2000, but has been growing slowly from 2000 to 2009. It even decreased by 0.1 percent every year since 2010. The chart in the lower right corner shows the distribution of US health care expenditures. The x-axis is the percent of population ranked by health care spending, and the y-axis is the percent of total health care spending. For example, the top 1 percent of patients account for 21.8 percent of total health care spending, and the top 50 percent of patients create 97.1 percent of health expenditures in 2009.

Health care spending is total payments from all sources, including direct payments from individuals and families, private insurance, Medicare, Medicaid, and miscellaneous other sources, to hospitals, physicians, other providers (including dental care), and pharmacies. Health insurance premiums are not included.

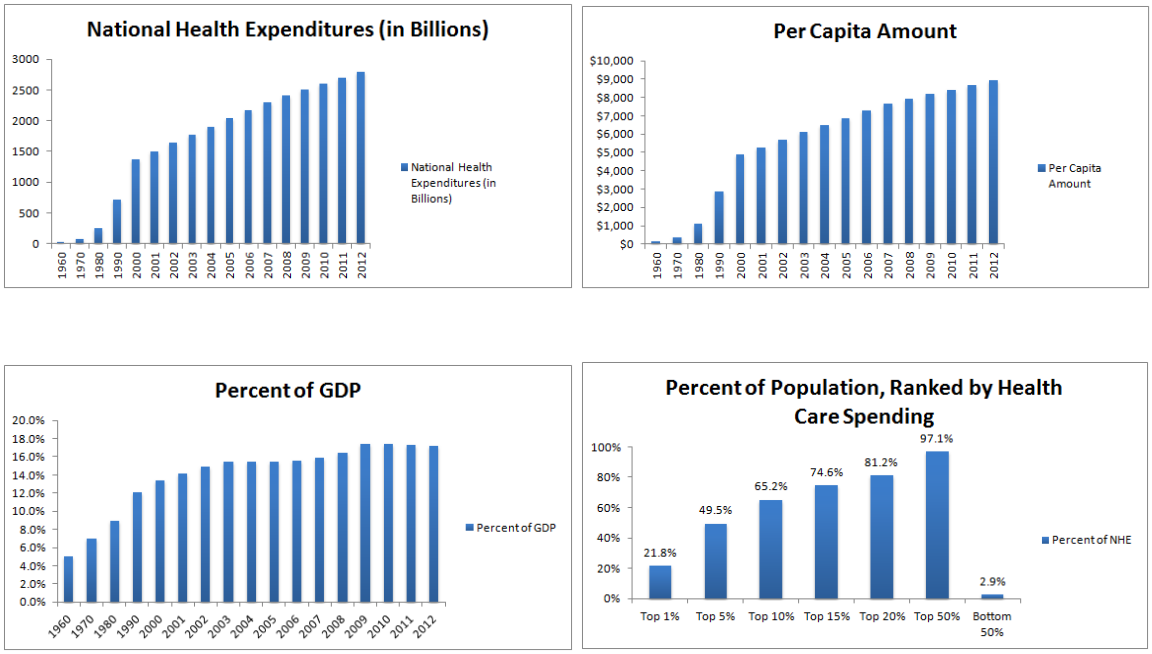


Figure 6.1: National Health Expenditures

Source: Centers for Medicare & Medicaid Services, National Health Expenditure Data

In current fee-for-service model, fee-for-service refers to the payment model where

health care providers receive a fee for each service such as an office visit, test, procedure, or other health care service. Payments are issued retrospectively, after the services are provided. Fee-for-service is the dominant physician payment method in the United States and Laugesen believes that [59]this mechanism is the major driver of high health care costs. Under this payment model, providers do not take risks since they get reimbursement for the services they provide. Therefore, it provides incentive for the provider to increase volume whether appropriate or not, without any risk. It also encourages duplication, discourages care coordination, and promotes inefficiency in the health care delivery system [69]. Similarly, when patients are shielded from paying (cost-sharing) by health insurance coverage, they are incentivized to welcome any medical service that might do some good. A variety of reform efforts have been attempted, recommended, or initiated to reduce its influence such as moving towards bundled payments and capitation. Despite all the criticism fee-for-service model gets, Ginsburg [48] believes that the core method of payment to many physicians for the services they provide is likely to remain fee-for-service, therefore it is critical to address the current shortcomings in the Medicare physician fee schedule.

Another payment approach is called capitation or global capitation. Under global capitation, provider is paid a set amount of money per patient per month, independently of the costs actually incurred. It usually pays a single health care organization. Global capitation is currently used by private HMOs, including in publicly financed products like Medicare Advantage plans and Medicaid managed care plans. If a patient insured through a globally capitated plan uses services that cost less than the amount paid to the provider organization, the organization keeps the leftover funds as profit. To ensure providers do not withhold needed care, globally capitated providers often have to report on quality and utilization measures, which can be linked to performance bonuses or publicly reported [10]. The amount of the global

payment can be based on normative standards (e.g., the average risk-adjusted payment for the population in the community) or based on historical spending for the population cared for by the capitated organization, trended forward. Global capitation encourages providers to improve efficiency of care, reduce unnecessary care and bring spending under control. In addition, unlike episode-based payments, capitation provides incentives to keep patients healthy, not just to limit the costs of episodes when they occur [4]. Physicians can also provide services to patients in various ways such as phone consultations that can improve efficiency but are difficult to incorporate into fee-for-service models. However, providers have to bear full risk beyond capitated payment.

Many of the payment plans used to pay health care providers in today's environment could be classified as bundled-services arrangements [26] or episode-based payments. From the perspective of CMS and other private payers, the goal of bundled payments is to improve care coordination and reduce the use of duplicative or unnecessary services [63]. The payments are designed to achieve this goal by creating financial incentives for hospitals and affiliated providers to keep the costs of surgical and acute care episodes below certain amount. Bundled payment introduces an incentive for providers to select lower-price treatment regimens from among those deemed equally appropriate - an incentive not present in the current fee-for-service system. A bundled-services payment plan has two key features. First, payments to the provider are not necessarily related to the list of specific services provided the patient and identified in the UB-92 or the CMS-1500. Instead payment is grouped into a mutually exclusive set of services categories. For example, hospitals are paid by some health care plans on a per-diem or per-case payment rate. Both are examples of bundled services payment. Second, bundled-services arrangements have a fixed fee specified per unit of service. For example, in the per diem arrangement, revenue

from treating a patient would be equal to the length of stay times the negotiated per-diem rate. Medicare has developed bundled-services payment plans for most health care providers. Medicare pays hospitals for inpatient care on a bundled-services unit basis referred to as PPS (Prospective Payment System). Medicare officially launched PPS on October 1, 1983. All participating in the Medicare program are required to participate in PPS, except those excluded by statute. PPS provides payment for all hospital non-physician services provided to hospital inpatients. This payment also covers services provided by outside suppliers, such as laboratory or radiology units. The basis of PPS payment is the DRG system developed by Yale University. The DRG system takes all possible diagnoses from the International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) system and classifies them into 25 major diagnostic categories based on organ systems. These 25 categories are further broken down into 559 distinct medically meaningful groupings or DRGs. Medicare contends that the resources required to treat a given DRG entity should be similar for all patients within a DRG category.

For hospitals that have higher costs across procedures might face substantial financial burden with bundled payments. On the other hand, those hospitals that have low costs of services would find bundled payment favorable to keep financially viable. Miller et al. [63] use multiple linear regression to describe hospital-level variation in Medicare payments for inpatient surgery. Their findings of the existence of wide variation in payments imply opportunities for substantial savings for CMS and other payers. However, the potential savings will depend strongly on the procedures and services selected for bundled payment programs. They also suggest the possible mixed effects of bundled payments on providers. For example, hospitals that are currently low cost for one or more of the services potentially covered by bundled payments, the payments may equal or exceed current fee-for-service reimbursements.

In this case, the financial incentive would be weak for these providers to focus on continued improvements in care coordination and cost efficiency.

In implementation of bundled payment, policy makers need to set base payment rates for episodes of care and update the rates over time to reflect changes in the costs of delivering care and the components of care. However, Rosen et al. [66] conduct analysis on 2003 and 2007 US commercial claims data showing that spending growth to be highly skewed across episodes. 10 percent of episodes accounted for 82.5 percent of spending growth, and within-episode spending growth ranged from -75 percent to 323 percent. Therefore, instead of updating the reimbursement rates uniformly across episodes of care, new approaches need to be developed to address variations in spending growth.

Consumer-directed health plans (CDHPs) , also known as high-deductible health plans (HDHPs), are relatively new in origin and became viable alternatives with the passage of the 2003 Medicare Prescription Drug, Improvement, and Modernization Act. They combine two elements. The first element is a health plan, usually a PPO, that has a high deductible (such as $5,000) and low premiums. The second element is a special "savings account" that is used to pay medical bills before the deductible has been met. The health savings account, similar to an individual retirement account (IRA), lets people put aside untaxed wages that they may use to cover their out-of-pocket medical expenses. Some employers contribute to employees' accounts as a benefit. The primary objective of these plans is to increase the involvement of patients in selecting cost-effective health care services. In 2009, the median deductible for individual coverage in a high deductible health plan was $1600, compared with $400 for traditional deductible health plans [20]. For 2011, the Internal Revenue Services (IRS) has set the minimum HDHP deductibles at $1200 for individual coverage and $2400 for family coverage. Most of the large health plans provide CDHP options,

but current enrollment in CDHPs still remains small. There is mixed effect of HDHPs on the utilization of care among patients enrolled in these plans. Wharam [85] analyze emergency department (ED) visits and hospitalizations over two years among enrollees insured in high-deductible plans through small employers in Massachusetts. They find that people of high socioeconomic status enrolled in high-deductible health plans did cut their use of emergency department visits for lower-severity conditions by 15-20%, with appropriate use of emergency care for serious conditions unchanged. However, people of low socioeconomic status experienced 25-30 percent reductions in high-severity ED visits over both years, while hospitalizations declined by 23 percent in year 1 but rose again in year 2, suggesting that delayed care led to even more serious illness requiring hospitalization. They suggest that policy makers and employers should consider proactive strategies to educate high-deductible plan members about their benefit structures or identify members at higher risk of avoiding needed care. They should also consider implementing means-based deductibles. Kozhimannil et al. [56] found similar disparities according to gender, with men enrolled in high-deductible plans more likely than women to forego needed care,which implies that clinicians caring for patients with HDHPs should be aware of sex differences in response to benefit design.

6.1.2 The Delivery System

In the healthcare delivery system, health services can be classified into three types: primary care, secondary care, and tertiary care. Primary care is preventive and wellness care, provided by physicians, or nurse practitioners. The goal of primary care is to decrease the health risk of individuals and the community. Usually, the physician or the nurse practitioner is the first person a patient would reach out to regarding any health issue. Secondary care is provided by a hospital or specialist,

referred by the primary care provider, that has more specialized knowledge, skill, or equipment. Tertiary care refers to more specialized and advanced consultative heath care, such as cancer management and plastic surgery, referred by primary care provider or secondary care provider.

The healthcare delivery system in the United States has been highly complicated and costly, and an Integrated Delivery System (IDS) has been attractive to all players in the system. IDS is a fairly broad concept. One definition of IDS is that "It is a network of health care providers and organizations which provides or arranges to provide a coordinated continuum of services to a defined population and is willing to be held clinically and fiscally accountable for the clinical outcomes and health status of the population served" [83].It is believed that higher level of integration will yield a more efficient healthcare delivery system. A variety of reform efforts under the Patient Protection and Affordable Care Act have been attempted, recommended, or initiated to reduce costs and improve quality of care. These include Partnership for Patients, Bundled Payments for Care Improvement, Comprehensive Primary Care Initiative, Accountable Care Organization Models, etc. Among them, Accountable Care Organizations, or ACOs, are groups of doctors, hospitals, and other health care providers, who come together voluntarily to give coordinated high quality care to the Medicare patients they serve. ACOs can be considered as a higher-level system based on IDSs. The goal of ACOs is to avoid unnecessary spending such as repeated diagnostic tests. ACOs also take responsibility of keeping the patients out of the hospital and emergency room and helping patients manage their conditions by getting them to take their medications appropriately and coming back for needed appointments. ACO is set up by providers, not an insurance company, so the doctors and hospitals can work within their own framework as long as they are meeting the 33 quality measures and outcomes agreed to the contract with the insurer. Doctors

and hospitals in ACOs are given a spending target for patient care with requirement on health care quality standards. They can keep some of the savings if they spend less than the budget, but will be penalized if they spend too much. There are 626 ACOs in the United States as of May 2014, both in public and private sectors.

Hwang et.al [54] review 25 related articles and manuscripts, and conclude an association between increased integration in healthcare delivery and an increase in the quality of care, in terms of clinical effectiveness, length of stay, medication errors, and number of office visits. As of cost reduction, some studies show that IDSs are associated with lower cost of care with level of service utilization being used as a proxy measure for cost of care, while other studies show no significant relationship between care integration/coordination and cost reduction.

6.2 Introduction to Robust Optimization

Traditional methods to address uncertainty in optimization are stochastic programming and dynamic programming. Stochastic programming assumes the parameters are uncertain but their distributions are known. A stochastic programming problem minimizes (or maximizes) the expected objective value over all possible scenarios caused by the uncertainties that follow certain distributions. Dynamic programming deals with multi-stage decision making, in which the value function is maximized such that the decision is optimal at all time periods (also called the optimal policy). However, stochastic programming and dynamic programming have two major drawbacks. First, the probability distribution of uncertainty is hard to be estimated accurately. Further, the size of the problem grows exponentially as the number of scenarios (for SP) or the number of states/time periods (for DP) increases, which makes problem solving very difficult or intractable.

Robust optimization (RO) comes in as another approach to address data uncertainty, in which uncertainty is not stochastic, but deterministic and set-based. Instead of minimize/maximize expected objective value as in SP, RO considers the worst-case scenario and minimize the maximum objective value or maximize the minimum objective value over the uncertainty set. Therefore, the optimal solution of RO is immune to data uncertainty since it guarantees that the optimal solution is feasible and efficient for any realization of the uncertainty set. Another advantage of RO is its tractability. Bertsimas et. al [13] state that "many well-known classes of optimization problems, including LP, QCQP, SOCP, SDP, and some discrete problems as well, have an RO formulation that is tractable".

A simple example of RO is as follows. Consider the linear programming problem:

$$\begin{aligned} \max \quad & c'x \\ \text{s.t.} \quad & Ax \leq b \\ & x \geq 0 \end{aligned} \tag{6.1}$$

Suppose A is the uncertain parameter and belongs to the uncertainty set U, then the robust counterpart of Problem(6.1) is

$$\begin{aligned} \max \quad & c'x \\ \text{s.t.} \quad & a_i'x \leq b_i, \quad \forall i, \forall a_i \in U \\ & x \geq 0 \end{aligned} \tag{6.2}$$

or equivalently

$$\begin{aligned} \max \quad & c'x \\ \text{s.t.} \quad & \max_{a_i \in U} a_i'x \leq b_i, \quad \forall i \\ & x \geq 0 \end{aligned} \tag{6.3}$$

where a_i is the ith column of matrix A'. The uncertainty can be modeled in different

ways. Soyster [73] considers the linear programming problem where the uncertainty belongs to column-wise convex sets. This method is too conservative since the optimal solution requires all uncertain parameters to be equal to their worst-case values. A less conservative approach proposed by Ben-Tal and Nemirovski [11] considers linear programming problem with ellipsoidal uncertainty sets, and the robust counterpart problem is turned into tractable conic quadratic problem. One drawback of this method is its nonlinear model structure with higher computational demands. Bertsimas and Sim [15] define uncertainty set as a polyhedron, which consists of range forecasts for each uncertain parameter. The robust counterpart of the linear problem is proved to be linear as well. In addition, they introduce a constraint called "budget of uncertainty", to control the conservatism at the will of decision maker. Suppose the uncertain parameter a_{ij} belongs to a symmetric, bounded interval $[\bar{a_{ij}} - \hat{a_{ij}}]$ where $\bar{a_{ij}}$ is the point forecast of a_{ij} and $\hat{a_{ij}}$ is the deviation from the nominal value. Define the scaled deviation y_{ij} as

$$y_{ij} = \frac{a_{ij} - \bar{a_{ij}}}{\hat{a_{ij}}}, \forall i, j \tag{6.4}$$

such that $a_{ij} = \bar{a_{ij}} + \hat{a_{ij}} y_{ij}$. The scaled deviation y belongs to the set:

$$Y = \left\{ y | \sum_{j=1}^{n} |y_{ij}| \leq \Gamma_i, \forall i, |y_{ij}| \leq 1, \forall i, j \right\}, \tag{6.5}$$

where $\Gamma \in [0, n]$ is the budge of uncertainty, which determines the number of a_{ij} that can be deviated from the nominal value $\bar{a_{ij}}$. If $\Gamma = 0$, then all parameters are certain. If $\Gamma = n$, then all parameters are uncertain. If $0 < \Gamma < n$, then the decision maker can control the level of conservatism while protect against parameter uncertainty to some degree. By incorporating the budge of uncertainty constraint, Problem (6.1)

becomes:

$$\begin{aligned} \max \quad & c'x \\ \text{s.t.} \quad & \bar{a_i}'x + \max_{y_i \in Y} \sum_{j=1}^{n} \hat{a_{ij}}' x_j y_{ij} \leq b_i, \quad \forall i \\ & x \geq 0. \end{aligned} \tag{6.6}$$

The robust problem (6.6) can be reformulated as a linear programming problem [16]:

$$\begin{aligned} \max \quad & c'x \\ \text{s.t.} \quad & \bar{a_i}'x - \Gamma_i p_i - \sum_{j=1}^{n} q_{ij} \leq b_i, \quad \forall i \\ & p_i + q_{ij} \leq \hat{a_{ij}} z_j, \qquad \forall i, j \\ & -z_j \leq x_j \leq z_j, \qquad \forall j \\ & p_i, q_{ij} \geq 0, \qquad \forall i, j \\ & x \geq 0. \end{aligned} \tag{6.7}$$

The reader is referred to Bertsimas et. al [13], Gabrel et. al [46] and references therein for a comprehensive review on theories, applications and recent advances of robust optimization.

6.3 Applications of Robust Optimization in Healthcare Systems Engineering

6.3.1 Robust Regression and Prediction of Healthcare Costs

Robust Regression

Ordinary least squares (OLS) linear regression can be very sensitive to outliers due to various assumptions of the model such as residuals having constant variance and are normally iid (independent and identically distributed). Robust regression is an alternative estimation method which down-weight or ignore unusual data such that it can still provide useful information when some of the assumptions are violated. When the weights given to each observations are close to one, the results of robust regression would be the same as OLS estimates. There are different types of robust regression models when it comes to how they give less weights to observations that would otherwise influence regression line. Some commonly used estimators include M-estimators, bounded-influence estimators, MM-estimators, and L_1-regression estimators. Robust regression can be implemented in SAS ("ROBUSTREG" procedure) and R ("robust" package).

M-Estimation: "M"-Estimation gets its name since it is considered as a generalization of the Maximum Likelihood Estimation method. It is first introduced by Peter J. Huber in 1964. M-estimation deals with outliers in the way that gives less weights to observations with large residuals [45]. An objective function f and a weight function w are involved in the estimation process. Consider a linear model

$$y_i = x_i^T \beta + \epsilon_i, \quad i = 1, \ldots, n, \tag{6.8}$$

then the residual would be

$$e_i = y_i - \hat{y}_i = y_i - x_i^T \hat{\beta}. \tag{6.9}$$

In regression estimation, the estimates $\hat{\beta}$ are obtained by minimizing the objective function

$$\sum_{i=1}^{n} f(e_i) = \sum_{i=1}^{n} f(y_i - x_i^T \hat{\beta}). \tag{6.10}$$

Intuitively, the minimization problem 6.10 can be solved by differentiating with respect to $\hat{\beta}$ and setting it to 0:

$$\frac{\partial}{\partial b} \sum_{i=1}^{n} f(y_i - x_i^T \hat{\beta}) = 0. \tag{6.11}$$

If we define the weight function w as the derivative of f with respect to $\hat{\beta}$, then equation 6.11 becomes

$$\sum_{i=1}^{n} w_i (y_i - x_i^T \hat{\beta}) x_i^T = 0. \tag{6.12}$$

Since weights w, coefficients β, and residuals e depend on each other, the minimization problem 6.10 can be solved by an iterative method until the estimated coefficients converge. For the traditional OLS method, the objective function $f(e) = e^2$ represents sum of squared residuals, and the weight function $w(e) = 1$ since it gives equal weight to every observation. However, for M-estimation, one possible objective function (Huber Method) and its corresponding weight function are:

$$f(e) = \begin{cases} \frac{1}{2}e^2, & |e| \le c \\ c|e| - \frac{1}{2}c^2 & |e| > c \end{cases}, w(e) = \begin{cases} 1, & |e| \le c \\ c/|e| & |e| > c \end{cases} \tag{6.13}$$

Bounded-influence estimation: Although M-estimators are insensitive to out-

liers and have relatively high efficiency, they are not robust to high leverage points under certain circumstances. Therefore, bounded-influence (BI) estimators are used to deal with high leverage points. One example is the least-trimmed squares (LTS) estimator, in which the estimates $\hat{\beta}$ are obtained by minimizing the sum of the smallest k of squared residuals. However, how to choose k can be tricky and they sometimes provide unreasonable results. Therefore, one estimation method which combines M-estimation and Bounded-Influence estimation is called MM-estimation, where the BI estimator provide starting points for M-estimation. This method takes advantage of M-estimation's high efficiency and BI estimation's high breakdown point.

L_1 **regression:** It is the most commonly used method of robust regression. As a special case of quantile regression, L_1 regression minimizes the sum of absolute residuals:

$$\hat{\beta} = \arg\min \frac{1}{n} \sum_{i=1}^{n} |y_i - x_i^T \beta|. \tag{6.14}$$

Compared with OLS method, L_1 gives much less weight to observations with large residuals. Instead of estimating mean of response variable y at $x^T\beta$, L_1 regression estimates median of y at $x^T\beta$.

Note that here we only consider robustness to outliers, not other types of robustness such as model misspecification. Minimax robust designs for misspecified regression models is proposed and illustrated by Heo et.al [53]and Shi et.al [71]. Robust regression could possibly be confused with Robust optimization in statistical estimation [47] [24], which assumes the coefficient matrices A and b are uncertain but bounded. For the purpose of healthcare costs prediction, we do not consider this type of robustness either.

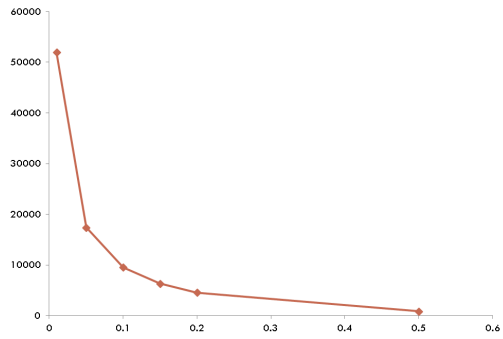
Application: Prediction of Healthcare Costs

Healthcare costs have been growing at unsustainable rates these years. National health spending grew 3.6% in 2013. Thus healthcare cost prediction plays an important role in improving accountability in care. To take a closer look at the distribution of health spending, the cut-off points representing the annual expenses per person in each percentile are shown in the table below. For instance, almost half of all health care spending (49.5%) was used to treat just 5% of the population, which included individuals with health expenses at or above $17,402 per year. We plot health spending versus percent of population and find that it obeys power-law distribution. The distribution of health spending is highly concentrated because the onset of disease is unpredictable and can require intensive technology and time to treat.

Figure 6.2: Cut-Off Points of NHE

Percentiles	Cost cut-off points
Top 1%	$51,951
Top 5%	$17,402
Top 10%	$9,570
Top 15%	$6,343
Top 20%	$4,586
Top 50%	$851

Figure 6.3: Distribution of NHE



Researchers have been using statistical models to better predict healthcare costs. Claims data and administrative data are two major utilized data sources. According to Duncan [34], there are two major types of models for healthcare costs prediction. One type is non-condition risk-based models. The other type is risk factor-based models. The non-condition risk-based models are mostly used for pricing, underwriting and candidates selection for care management programs. They use age/sex only, prior cost only, or combination of these two predictors. For the age/sex only

model, relative factor ratios are created and multiplied by baseline health costs to predict future cost of an individual or a group of individuals. For the prior cost only model, experience trend factor is multiplied by baseline PMPY (per member per year) when calculating subsequent year predicted cost. For the combination model, the expected cost is calculated as the weighted sum of prior cost and age/sex rating. The combination model produces more accurate predictions than individual models. The accuracy of all three models improves as the size of group increases. The prior cost model is more accurate when groups have outlier members who have extra claims. Since non-condition risk-based models use very limited information about individual risk factors, their predictions are less accurate than models incorporating additional individual risk factors. Risk factor-based models utilize factors such as medical condition related factors and life style related factors, besides age and gender. Common statistical methods that risk factor-based models are built on include regression, classification trees, and clustering algorithm.

Most literature use regression models to analyze and predict healthcare costs. In Gregori et. al's review [51], different models are designed to match the characteristics of healthcare costs. To address the high skewness of the distribution of healthcare costs, three methods are widely used: (1) OLS on transformed response variable, which may cause interpretation problems during back transformation; (2) Threshold model, which estimates the probabilities that the costs are greater than the median and the third quantile, thus does not give an estimate of the mean; (3) Generalized Linear Model (GLM), which transform the expectation instead of response variable itself, thus is a flexible approach. To address the mass at zero costs, the most popular methods are: (1) Add a positive constant to the costs, which performs poorly as the choice of the constant is tricky and it does not take into account the differences between "true" positive costs and "fake" positive costs; (2) Tobit model, which

introduces a new binary variable with value 1 when $c > 0$ and with value 0 when $c \leq 0$; (3) Mixed model, in which the conditional expectation is

$$E(c_i|x) = p(c_i > 0)E(c_i|c_i > 0). \tag{6.15}$$

Classification and clustering algorithms are also being used to predict healthcare cost buckets. Bertsimas et. al [12] utilize data mining techniques on claims data from over 800,000 insured individuals over three years. They divide the learning sample into five cost buckets and apply decision trees and clustering algorithms to predict the median dollar amount of healthcare costs. The baseline method is to simply use last 12 months of observation period as forecast of overall healthcare cost in the next period. They validate their models using over 200,000 out-of-sample members. The independent variables include diagnosis groups, procedure groups, drug groups, cost variables, age and gender etc. The performance measures are the hit ratio, the penalty error, and the absolute prediction error. They findings are: (1) classification tree algorithm does a bit better on lowest-cost buckets for the hit ratio and penalty error, but the clustering algorithm performs better on the higher-cost buckets; (2) the pattern of past cost data is a strong predictor of future costs; (3) medical information only contributes to accurate prediction of medical costs of high-cost members. To improve the classification trees used by Bertsimas, random forests can be utilized since random forests is "unexcelled in accuracy among current algorithms and runs efficiently on large data bases." Random forests is an ensemble model of classification trees. There are two sources of "randomness" in the model: (1)randomly sample N data points with replacement from original dataset as training set for tree growing; (2)m variables out of M input variables are randomly selected and are used to best split the node. Advantages of random forests include: (1)It does not overfit; (2)There is no need for cross-validation or a separate test set to get

an unbiased estimate of the test set error. It is estimated internally during the run. Random forests is particularly suitable for imbalanced data as healthcare cost data. No literature is found in healthcare costs prediction using random forests.

In addition to regression and classification/clustering methods, time series analysis is utilized by Vliet [82], who uses healthcare expenditure and insurance coverage data of about 35,000 members enrolled with the largest private health insurer in the Netherlands to estimate an upper bound on the proportion of variance in annual individual healthcare expenditures. The results show that at most 20 percent of the variance is predictable. Vliet's study considers four time series models: VC (Variance Components) model, AR (AutoRegressive) model, ARVC (AutoRegressive-Variance Components) model, and ARMA (AutoRegressive-Moving Averages) model.

Robust regression can be applied in prediction of healthcare costs due to the heavy-tailed distribution of healthcare costs. Literature in this application is relatively limited. Szpiro et.al [80] present a new Bayesian approach to model-robust linear regression which leads to uncertainty estimates with the same robustness properties as the Huber-White sandwich estimator. They study the relationship between average annual outpatient healthcare costs and age using data from Washington State Basic Health Plan. Robust regression is proved to perform well with high accuracy and low MAPE (Mean Absolute Percentage Error) [57] [62].

6.3.2 Robust Markov Decision Processes and Disease Management

Robust Markov Decision Process

Markov Decision Processes (MDP) is a powerful tool for sequential decision making under uncertainty. It is a higher-level decision tree model which addresses more complicated problems and utilizes more advances modeling techniques. As an extension of Markov chains, it differentiates with respect to multiple options in actions and rewards. It has been applied in many medical treatment decision making problems, such as epidemic control, drug infusion, kidney/liver transplantation, treatment of ischemic heart disease, etc [68]. There are two ways to classify MDP. It may be classified into discrete-time, where decisions are made at discrete time intervals or continuous-time MDP, where decisions can be made at any time. It could also be classified into finite-horizon or infinite-horizon MDP according to the time horizon in which the decisions are made [9].

All MDPs have five basic components: (1) T: the set of points in time when decisions are made (discrete interval vs. continuous interval); (2) S: the set of all possible states the system could be at; (3) A: the set of possible actions that the decision maker could take for each state; (4) $p_t(s'|s,a)$: the transition probability that action a in state s at time t will lead to state s' at time $t+1$; (5) $r_t(s,a)$: the expected immediate reward by taking action a at state s. A discount factor $\gamma \in [0,1]$ is often present in MDPS when future rewards are discounted over time. For instance, $\gamma = \dfrac{1}{1+r}$ if the discount rate is r.

The problem of MDPs is to find an optimal policy π that maximizes a measure of rewards, which is typically the expected discounted sum of rewards over a finite or

infinite horizon. We introduce briefly the problems and algorithms of discrete-time and continuous-time MDPs. The object function of discrete-time MDP is:

$$\max_{a_t \in A} \sum_{t=0}^{\infty} \gamma^t r_t(s, a) \tag{6.16}$$

where $0 \leq \gamma < 1$, and the maximum value is obtained when $a_t = \pi(s_t)$. Discrete MDPs can be solved by linear programming or dynamic programming. Two most commonly used variants of dynamic programming are value iteration and policy iteration. The value iteration method solves Bellman equations iteratively backwards in time:

$$y_T^*(s_T) = r_T(s_T) \qquad \forall s_T \in S, \tag{6.17}$$

$$y_t^*(s_t) = \max_{a \in A} \left\{ r_t(s_t, a) + \gamma \sum_{i \in S} p_t(i|s_t, a) y_{t+1}^*(i) \right\}, \quad t = 1, \cdots, T, \quad s_t \in S \tag{6.18}$$

where $y_T^*(s_T)$ denotes the terminal reward at time T when the state is s_T, and $y_t^*(s_t)$ represents the optimal value of total expected reward when the state at time t is s. The optimal policy is comprised of optimal actions $a_{s_t,t}^*$ given by equation (6.19).

$$a_{s_t,t}^* \in \arg\max_{a \in A_s} \left\{ r_t(s_t, a) + \gamma \sum_{i \in S} p_t(i|s_t, a) y_{t+1}^*(i) \right\}, \quad t = 1, \cdots, T \tag{6.19}$$

The policy iteration method is quite different since it initializes the process by choosing an arbitrary policy π. The next step is policy evaluation, where the total expected reward is calculated by solving a set of linear equations under the chosen arbitrary policy. Followed by policy evaluation is policy improvement, where a better policy is obtained and updated for each state s under the assumption that the arbitrary policy is used for the next step onwards. This process is repeated until π does not change any more or converges.

The formulation of the linear programming problem for MDP, where $y^*(s)$ for every state s are variables, is [52]:

$$\begin{aligned} \min \quad & \sum_s y^*(s) \\ \text{s.t.} \quad & y^*(s) - r_t(s_t, a) - \gamma \sum_{i \in S} p_t(i|s_t, a) y^*_{t+1}(i) \geq 0, \quad \forall s, a. \end{aligned} \tag{6.20}$$

The optimal solution of Problem (6.20) can be obtained by solving its dual program. For continuous-time MDP, the objective function becomes:

$$\max_{a(t) \in A} E[\int_0^\infty \gamma^t r(s(t), a(t)) dt]. \tag{6.21}$$

Continuous-time MDPs can be solved in similar algorithms as discrete-time MDPs, and are omitted here.

In practice, the transition probabilities and the reward parameters in the above MDP models and algorithms are either estimated from historical data or learned from experience. However, it is hard to get access to enough historical data or to accurately estimate the parameters, which may lead to significant difference between long-term performance of a strategy and the model's prediction [32]. Therefore, robust optimization comes in to incorporate the uncertainty in MDPs, thus called robust MDPs. Current literature on robust MDPs can be classified in two ways: classification based on type of objective function, and classification based on type of uncertainty source. With respect to the objective function, one type of robust MDPs maximizes the value function under worst case scenario, and the other type minimizes the maximum regret. With respect to uncertainty source, some studies focus on uncertain transition matrix P, while other studies focus on uncertain reward R. Wiesemann et.al [86] derive rectangular uncertainty sets for transition matrices P with pre-specified probability $1 - \beta$ from historical data. They assume that

the transition samples are not independent for each state-action pair, and consider both rectangular and non-rectangular uncertainty sets. The optimal policy that achieves maximum value under worst-case scenario with rectangular uncertainty set is obtained by tractable second-order conic programs. However, the robust MDP involving non-rectangular uncertainty set is intractable, and only approximate solutions can be obtained by semidefinite programming. Nilim and El Ghaoui [64] argue that polytope set is a poor representation of uncertain transition matrices and lead to over conservative robust solutions. They propose to model uncertainty via a Kullback-Leibler divergence bounds, and solve finite/infinite robust MDPs using the "robust dynamic programming" algorithm. The complexity of their algorithm is almost the same as the original Bellman recursion in dynamic programming, thus leading to no extra computational costs. Xu and Mannor [90] consider the trade-off between worst case performance and nominal performance over all models to address the over-conservatism of robust MDP models. Their algorithm computes the trade-off between robustness and performance by optimizing the weighted sum of the robustness criterion and the performance criterion. The algorithm is also applied to robust MDPs and it is tractable only for MDPs with uncertain reward parameters, not for uncertain transition matrices. Ahmed et.al [8] formulate the regret minimization problem over possible models of dependent transition and reward uncertainty, and approximate it as a Mixed Integer Linear Program. The regret for a policy π is defined as the difference between the value of optimal policy π^* and the value of policy π.

Application: Disease Management

Adherence of medication is a very important problem since nonadherence may lead to serious complications and hospitalization. Since there are probabilities of adherence

and decisions on how often/long a nurse makes a call are involved, Markov Decision Process is an option to facilitate the decision making process. Mason et.al [60] use several MDP models to determine the optimal time of intervention to improve medication adherence, and the MDPs are solved by backwards recursion. They consider the trade-off between maximum patient adherence and minimum cost of intervention. They find out that it is cost effective to implement Electronic Health Records-based active surveillance system for cardiovascular disease management program.

Consider the simple disease management problem where a nurse tracks the medication adherence of a patient recovering in acute care by weekly phone calls. Suppose the recovery time for the acute illness is three weeks. At the beginning of the first week, the nurse makes either a long call or a short call to the patient. After that, at the end of each week (including the first week), the nurse calls the patient to check if the patient adhered medication in the previous week or not. And the nurse makes the decision, based on the result of previous week, on whether to make a long call or a short call to maintain/improve adherence. We model this problem as a discrete-time finite MDP, in which time horizon T is 3 weeks; S is either 1 (adherence) or 0 (nonadherence); A is the action the nurse could take which is either making a long call or a short call; transition probability matrix P describes the probabilities of adherence after long call and short call; and the reward R is the total cost including cost of nurse intervention and the cost of hospitalization due to nonadherence. The objective is to find the optimal policy such that the total cost is minimized:

$$y_t^*(s_t) = \min_{a \in A} \left\{ r_t(s_t, a) + \gamma \sum_{i \in S} p_t(i|s_t, a) y_{t+1}^*(i) \right\}, \quad t = 1, \cdots, T, \quad s_t \in S. \tag{6.22}$$

The transition probabilities can be obtained by observing historical data. However, the estimations are subject to error. Therefore, we could construct uncertainty set for the transition matrix, for instance, a state-wise uncertainty set. Suppose the

transition matrix P belongs to the uncertainty set U, then the objective function of the robust MDP problem becomes:

$$\min_{\pi} \max_{P \in U} \left\{ r_t(s_t, a) + \gamma \sum_{i \in S} p_t(i|s_t, a) y^*_{t+1}(i) \right\}, \quad t = 1, \cdots, T, \quad s_t \in S. \tag{6.23}$$

The problem can be solved by the robust dynamic programming algorithm (See Nilim and El Ghaoui [64] for details of the algorithm). For this particular problem, robust dynamic programming algorithm would be sufficient due to the small size of the problem. If the size of the problem increases as the number of time periods goes to a large number, then robust dynamic programming algorithm might not be as efficient as it is now due to the "curse of dimensionality". Tamar, Mannor and Xu [81] propose a robust approximate dynamic programming method based on a projected fixed point equation to approximately solve large scale robust MDPs. They demonstrate the effectiveness of their method through simulation of an option pricing problem.

6.3.3 Robust Linear Programming and IMRT Fluence Map Optimization

Application: IMRT Fluence Map Optimization

Intensity modulated radiation therapy (IMRT) is one of the advanced cancer treatments that uses external radiation beams to irradiate tumors. The goal of the treatment is to deliver a prescribed amount of radiation to precisely conform to the 3D shape of the tumor, while sparing the surrounding critical organs and normal tissues. There are three major optimization problems in IMRT: beam-angle optimization, fluence map optimization and intensity delivery optimization. Among them, fluence

map optimization aims to determine the optimal intensity profile of each beam given a set of predefined radiation beams. The general linear formulations of the fluence map optimization problem is described in Yih [92]. Let (x, y, z) denote a point in the three-dimensional treatment volume including target (T),critical organs(O), and normal tissues(S); $(\bar{m}, k)$ denotes the beamlet k of beam $\bar{m}$, $k = 1, ..., t$; $v_{(x,y,z,,\bar{m},k)}$ denotes the dose contribution from beamlet $(\bar{m}, k)$. The fluence map optimization problem can be formulated as a linear programming problem with fluence values $w_{\bar{m},k}$ as decision variables:

$$\begin{aligned} \min \quad & f(V_{(x,y,z)}) \\ \text{s.t.} \quad & \sum_{\bar{m}\in\bar{M}} \sum_{k=1}^{t} w_{\bar{m},k} \cdot v_{(x,y,z,,\bar{m},k)} \leq V_{(x,y,z)} \\ & w_{\bar{m},k} \geq 0, \quad \forall \bar{m} \in \bar{M}, k = 1, ..., t \end{aligned}$$

The objective function and constraints can be defined differently. An example of the objective function is to minimize the sum of total deviation between the delivered dose and prescribed dose at all points:

$$f(V_{(x,y,z)}) = \rho_T f(V_T) + \rho_O f(V_O) + \rho_S f(V_S), \tag{6.24}$$

where $f(V_T)$,$f(V_O)$, and $f(V_S)$ are linear functions of decision variable $w_{\bar{m},k}$, and ρ_T,ρ_O and ρ_S are weighting factors. An example of the constraints limits doses on points within the tumor target in the interval $[T_l, T_u]$: $T_l \leq V_T \leq T_u$. Another important constraint is the dose volume constraint, which limits the RELATIVE volume of a structure that receives more or less than a particular threshold [58].

Various types of uncertainties exist in the IMRT process such as periodic breathing and cardiac motion within a treatment, as well as changes in intra-abdominal pressure and weight changes over the course of treatment [25]. Common methods

to address include adding safety margins to targets or using a motion probability density function to represent dose distribution. Bortfeld et.al [18] proposes a robust optimization framework to incorporate the uncertainty in breathing motion for lung tumors. Their robust solution requires about 11% less dose to the healthy tissue than the margin solution (worst-case scenario), while providing the same level of protection against breathing uncertainty. The patient's breathing motion is modeled using a probability mass function(PMF). The breathing motion PMF for a patient in a particular fraction specifies the proportion of time the patient spends in each of a finite number of breathing motion states during that fraction. The uncertainty set is then a set of breathing motion PMF's that could be realized during treatment. Let $\mathcal{P}$ be the set of all PMF's on finite set X:

$$\mathcal{P} = \{\mathbf{p} \in \mathbb{R}^{|X|} | \forall x \in X, p(x) \geq 0; \sum_{x \in X} p(x) = 1\}. \tag{6.25}$$

Then the PMF uncertainty set would be a bounded polyhedron defined by a lower bound vector and an upper bound vector:

$$P = \{\mathbf{p} \in \mathcal{P} | \forall x \in X, \bar{p}(x) - p^{-}(x) \leq p(x) \leq \bar{p}(x) + p^{+}(x)\}. \tag{6.26}$$

They define the nominal problem incorporating motion as:

$$\begin{aligned} \min_{w} \quad & \sum_{v \in \mathcal{V}} \sum_{b \in \mathcal{B}} \sum_{x \in X} d_{v,x,b} p(x) w_b \\ \text{s.t.} \quad & \sum_{b \in \mathcal{B}} \sum_{x \in X} d_{v,x,b} p(x) w_b \geq t_v, \forall v \in \mathcal{V} \\ & w_b \geq 0, \forall b \in \mathcal{B}, \end{aligned} \tag{6.27}$$

where $d_{v,x,b}$ describes the dose delivered to voxel v, when the anatomy is in breathing phase x per unit intensity of beamlet b. Therefore, $\sum_{x \in X} \sum_{b \in \mathcal{B}} d_{v,x,b} p(x) w_b$ is the sum

of the doses to voxel v delivered under each breathing motion state weighted by the corresponding proportions of time spent in those states. By adding the PMF uncertainty set, the robust counterpart of Problem (6.27) would be:

$$\begin{aligned} \min_{w} \quad & \sum_{v\in\mathcal{V}}\sum_{b\in\mathcal{B}}\sum_{x\in X} d_{v,x,b}p(x)w_b \\ \text{s.t.} \quad & \sum_{b\in\mathcal{B}}\sum_{x\in X} d_{v,x,b}\tilde{p}(x)w_b \geq t_v, \forall v\in\mathcal{V}, \forall \tilde{p}\in P \\ & w_b \geq 0, \forall b\in\mathcal{B}. \end{aligned} \tag{6.28}$$

Problem (6.29) is equivalent to the tractable linear program:

$$\begin{aligned} \min_{w} \quad & \sum_{v\in\mathcal{V}}\sum_{b\in\mathcal{B}}\sum_{x\in X} d_{v,x,b}p(x)w_b \\ \text{s.t.} \quad & \sum_{b\in\mathcal{B}}\sum_{x\in X} d_{v,x,b}p(x)w_b - \sum_{b}\sum_{x\in X} d_{v,x,b}p^-(x)w_b + \sum_{x\in X} p^-(x)y_v - \sum_{x\in X} z_{v,x} \geq t_v, \forall v\in\mathcal{V} \\ & \sum_{b} d_{v,x,b}(p^-(x)+p^+(x))w_b \geq (p^-(x)+p^+(x))y_v - z_{v,x}, \quad \forall v\in\mathcal{V}, \forall x\in X \\ & w_b \geq 0, z_{v,x}\geq 0, \quad \forall b\in\mathcal{B}, \forall v\in\mathcal{V}, \forall x\in X. \end{aligned} \tag{6.29}$$

An adaptive robust optimization method is proposed by Chan et.al [23] to continuously incorporate observed breathing motion PMF after treatment starts into current uncertainty set and generate new uncertainty set.

6.3.4 Robust Mixed Integer Programming and Operating Room Planning

Application: Operating Room Planning

Cost reduction has become more and more crucial for hospitals under the Affordable Care Act. As one of the most expensive resources in the hospitals, operating rooms

receive massive attention regarding their planning and scheduling to improve operation efficiency and optimize financial returns. We refer to [22] for a comprehensive literature review of operating room planning and scheduling. Among all operating room planning and scheduling problems, the Surgical Case Assignment Problem (S-CAP) refers to the problem of assigning patients to a given set of operating room (OR) blocks over the planning horizon. There are two major sources of uncertainties existed in OR planning: emergency arrivals and surgery durations. We focus on OR planning under uncertain surgery durations. Simulation, stochastic optimization, and robust optimization can be utilized to address this type of uncertainty.

Addis et. al. [7] propose a robust mixed integer programming (MIP) model for the SCAP with uncertain surgery duration, aiming to minimize a penalty function associated with waiting time, urgency and tardiness of patients due to delay in serving patients, without generating scenarios. The deterministic formulation of the problem is:

$$
\begin{aligned}
\min \quad & \sum_{i\in I}\sum_{j\in J}\sum_{t\in T_j}([tu_i]+[(w_i+t-l_i)^+]u_i)x_{ij}^t+ \\
& \sum_{i\in I}([w_i+|T|+1]u_i+[(w_i+|T|+1-l_i)^+]u_i)(1-\sum_{j\in J}\sum_{t\in T_j}x_{ij}^t) \\
\text{s.t.} \quad & \sum_{j\in J}\sum_{t\in T_j}x_{ij}^t\le 1 \quad \forall i\in I \\
& \sum_{i\in I}s_i x_{ij}^t\le c_j+o_j^t \quad \forall j\in J,\forall t\in T_j \\
& o_j^t\le \delta v_j^t \quad \forall j\in J,\forall t\in T_j \\
& \sum_{j\in J}\sum_{t\in T_i}v_j^t\le \Delta \\
& x_{ij}^t\in\{0,1\} \quad \forall j\in J,\forall t\in T_j
\end{aligned}
\tag{6.30}
$$

where decision variable x_{ij}^t is a binary variable that equals to 1 if patient i is assigned to block j in time $t\in T_j$. To address the uncertain surgery duration s_i, an uncertainty

set which consists of intervals $[\bar{s}_i - \hat{s}_i, \bar{s}_i + \hat{s}_i]$ is constructed by setting $\bar{s}_i$ to the mean of the distribution of s_i, and $\hat{s}_i$ to the standard deviation of the distribution of s_i. A budget of uncertainty constraint is also incorporated. The robust counterpart of Problem (6.30) is formulated by replacing the constraint $\sum_{i \in I} s_i x_{ij}^t \leq c_j + o_j^t$ with the following constraints:

$$\begin{aligned} \sum_{i \in I} \bar{s}_i x_{ij}^t + \Gamma p^{jt} + \sum_{i \in I} q_i^{jt} \leq c_j + o_j^t, \qquad & \forall j \in J, \forall t \in T_j \\ p^{jt} + q_i^{jt} \geq \hat{s}_i x_{ij}^t, \qquad & \forall i \in I, \forall j \in J \\ p^{jt}, q_i^{jt} \geq 0, \qquad & \forall i \in I, \forall j \in J, \forall t \in T_j \end{aligned} \tag{6.31}$$

where Γ is the budget of uncertainty parameter which is the maximum number of patients with surgery duration reaching the upper bound of nominal duration. The optimal solution of the robust problem is obtained by applying the approach in Bertsimas and Sim [15]. Their experiments demonstrate a reduction in the number of surgery cancelation compared to the deterministic model. Also, by properly tuning the value of Γ, a reasonable trade-off between utilization rate and quality of service can be achieved.

Denton et. al. [33] consider the situation of multiple ORs under certainty. Two decisions are made by the models: how many ORs to open on a given day, and allocation of surgeries to ORs. They first propose a two-stage stochastic linear program with binary decisions in the first stage and simple recourse in the second stage. Followed by the stochastic programming problem is its robust counterpart, in which the objective is to minimize the maximum total fixed and variable costs associated with an uncertainty set for surgery durations. The robust counterpart is reformulated as a mixed integer (linear) program, and the numerical results show that it is a fast and effective heuristic for computing near-optimal solutions to the stochastic recourse problem. Since they also incorporate the budget of uncertainty constraint

in the robust problem, a sensitivity analysis is performed with respect to the choice of Γ, which is the maximum number of surgery blocks that reach their upper bounds. Intuitively, as Γ increases, more ORs tend to be opened since more surgeries reach their upper bounds in the worst case. It is illustrated by the experiments that for the 15-surgery block instances, the solution quality reaches its peak when Γ is in the range of 2-4.

6.3.5 Conclusions

Decision makers in the healthcare systems need to make informed decisions with imperfect information on a daily basis. As an approach to address uncertainty, robust optimization has been utilized in a variety of applications in healthcare systems engineering. In this chapter, we provide an overview of the healthcare systems, including payment systems and delivery systems in the United States. We survey applications of robust optimization, particularly in prediction of healthcare costs, disease management, IMRT fluence map optimization, and operating room planning, among others. Further, we propose a robust Markov Decision Processes model for the problem of disease management.

Chapter 7

Conclusions and Future Work

In this dissertation, we have addressed several problems of current healthcare financing systems in the United States. In chapter 2, we proposed a highly tractable approach to incorporate parameter uncertainty into calculation of risk scores and money transfers in the commercial risk adjustment program. The worst-regret is minimized to ensure more fair payments to health insurers under parameter estimation errors. The problem is solved via linear programming and empirical results show small changes in robust risk scores but large fluctuations in money transfers between insurers. In chapter 3, we investigated major attributes of plan premiums, and the regression results suggest that "company" is the most important attribute. We proposed a simple algorithm to identify inefficient plans and assigned three grade letters to health insurers in Pennsylvania and Massachusetts. We also have made recommendations on how much the "dominated" should decrease their premiums in order to be efficient in our framework. Both chapter 4 and chapter 5 analyze issues related to Medicare, where chapter 4 identifies main factors driving Medicare beneficiaries' choices of prescription drug plans between MAPD and PDP. Since the characteristics of these two groups of people are quite different, recommendation-

s are made to help policy makers better promoting their Part D plans. Similarly, chapter 5 analyzes the trends in Medicare services usage and reimbursement rate for all states. The correlation analysis disagrees with the concern that physicians tend to upcode on purpose. We turned our attention back to robust optimization in chapter 6, where we surveyed applications of robust regression, robust Markov Decision Processes, robust linear programming and robust mixed integer programming in healthcare systems engineering. We also proposed a robust MDP model for disease management problem where a nurse tracks medication adherence of a patient recovering from acute care.

There remain some areas of interest to us for future work. First, a large number of previously uninsured nonelderly people entered the system due to the Affordable Care Act of 2010. Incorporating learning in the risk adjustment models is needed, since more information will become available over time for the newly insured when they submit claims. In addition, application of analytics is increasingly prevalent in the healthcare industry. There is a demand of developing more analytical models particularly for problems in healthcare financing systems. With the aid of the highly quantitative and analytical tools, policy makers can make more dynamic and high-quality decisions that will most benefit the community they serve. In addition, because of the uncertain environment of healthcare systems, it is possible to extend applications of robust optimization to other problems in healthcare systems engineering, if suitable.

Bibliography

[1] Risk assessment and risk adjustment. *American Academy of Actuaries Issue Brief*, May 2010.

[2] Risk adjustment implementation issues. Technical report, Centers for Medicare and Medicaid Services, 2011.

[3] Summary of the affordable care act. Technical report, Kaiser Family Foundation, 2011.

[4] Farewell to fee-for-service? a "real world" strategy for health care payment reform. Working Paper 8, UnitedHealth Center for Health Reform & Modernization, December 2012.

[5] Issues for risk adjustment in medicare advantage. Technical report, Medpac, 2012.

[6] National health expenditures 2012 highlights. Technical report, Centers for Medicare and Medicaid Services, 2013.

[7] B. Addis, G. Carello, and E. Tanfani. A robust optimization approach for the operating room planning problem with uncertain surgery duration. In *Proceedings of the International Conference on Health Care Systems Engineering*,

volume 61 of *Springer Proceedings in Mathematics & Statistics*, pages 175–189. Springer, 2014.

[8] A. Ahmed, P. Varakantham, Y. Adulyasak, and P. Jaillet. Regret based robust solutions for uncertain markov decision processes. In *Advances in Neural Information Processing Systems 26*, pages 881–889, 2013.

[9] O. Alagoz, H. Hsu, A. J. Schaefer, and M. S. Roberts. Markov decision processes: A tool for sequential decision making under uncertainty. *Medical Decision Making*, 30:474–483, 2010.

[10] P. C. Alguire. Understanding capitation. *Philadelphia: American College of Physicians*, 2012.

[11] A. Ben-Tal and A. Nemirovski. Robust solutions to uncertain programs. *Operations Research Letters*, 25:1–13, 1999.

[12] D. Bertsimas, M. V. Bjarnadottir, M. A. Kane, J. C. Kryder, R. Pandey, S. Vempala, and G. Wang. Algorithmic prediction of health-care costs. *Operations Research*, 56(6):1382–1392, Nov.-Dec. 2008.

[13] D. Bertsimas, D. B. Brown, and C. Caramanis. Theory and applications of robust optimization. *SIAM Review*, 53(3):464–501, 2011.

[14] D. Bertsimas and M. Sim. The price of robustness. *Operations Research*, 52(1):35–53, 2004.

[15] D. Bertsimas and M. Sim. The price of robustness. *Operations Research*, 52(1):35–53, 2004.

[16] D. Bertsimas and A. Thiele. *Tutorials in Operations Research*, chapter Robust and data-driven optimization: Modern decision-making under uncertainty, pages 95–122. INFORMS, 2006.

[17] L. J. Blumberg and J. Holahan. The affordable care act can survive low enrollment and adverse selection in the first year. Timely analysis of immediate health policy issues, Robert Wood Johnson Foundation, December 2013.

[18] T. Bortfeld, T. C. Chan, A. Trofimov, and J. N. Tsitsiklis. Robust mangement of motion uncertainty in intensity-modulated radiation therapy. *Operations Research*, 56(6):1461–1473, 2008.

[19] J. Brown, M. Duggan, I. Kuziemko, and W. Woolston. Will risk adjustment decrease health care costs? new evidence from the medicare advantage program. Technical report, National Bureau of Economic Research, 2011.

[20] U. S. Bureau and L. Statistics. Program Perspectives on High Deductible Health Plans. 3(4):4–7.

[21] U. S. C. Bureau. State & county quickfacts.

[22] B. Cardoen, E. Demeulemeester, and J. Belien. Operating room planning and scheduling: A literature review. *European Journal of Operational Research*, 201:921–932, 2010.

[23] T. C. Y. Chan and V. V. Misic. Adaptive and robust radiation therapy optimization for lung cancer. *European Journal of Operational Research*, 231(3):745–756, 2013.

[24] S. Chandrasekaran, G. Golub, M. Gu, and A. Sayed. Parameter estimation in the presence of data uncertainties. *SIAM J. Matrix Anal. Appl.*, 19(1):235–252, January 1998.

[25] M. Chu, Y. Zinchenko, S. G. Henderson, and M. B. Sharpe. Robust optimization for intensity modulated radiation therapy treatment planning under uncertainty. *Phys. Med. Biol.*, 50(23):5463–77, Dec 2005.

[26] W. O. Cleverley, J. O. Cleverley, and P. H. Song. *Essentials of Health Care Finance*. Jones & Bartlett Learning, 7 edition, 2010.

[27] R. R. Cline, M. M. Worley, S. W. Schondelmeyer, J. C. Schommer, T. A. Larson, D. L. Uden, and R. S. Hadsall. Pdp or ma-pd? medicare part d enrollment decisions in cms region 25. *Research in Social and Administrative Pharmacy*, 6:130–142, 2010.

[28] F. Cribari-Neto and A. Zeileis. Beta regression in r. *Journal of Statistical Software*, 34:1–24, 2010.

[29] D. M. Cutler and R. J. Zeckhauser. Adverse selection in health insurance. Frontiers in Health Policy Research, volume 1, MIT, January 1998.

[30] K. Davis, C. Schoen, and K. Stremikis. Mirror, mirror on the wall: How the performance of the u.s. health care system compares internationally, 2010 update. Technical report, The Commonwealth Fund, June 2010.

[31] R. Day and P. Nadash. New state insurance exchanges should follow the example of massachusetts by simplifying choices among health plans. *Health Affairs*, 31(5):982–989, 2012.

[32] E. Delage and S. Mannor. Percentile optimization for markov decision processes with parameter uncertainty. *Operations Research*, 58(1):203–213, 2010.

[33] B. T. Denton, A. J. Miller, H. J. Balasubramanian, and T. R. Huschka. Optimal allocation of surger blocks to operating rooms under uncertainty. *Operations Research*, 58(4):802–816, July-August 2010.

[34] I. Duncan. *Healthcare Risk Adjustment and Predictive Modeling*. ACTEX Publications,Inc., 2011.

[35] R. Ellis. *Risk adjustment in health care markets: concepts and applications*. Wiley, 2008.

[36] S. L. Ferrari and F. Cribari-Neto. Beta regression for modelling rates and proportions. *Journal of Applied Statistics*, 31(7):799–815, 2004.

[37] C. for Medicare & Medicaid Services. Ma state/county penetration.

[38] C. for Medicare & Medicaid Services. Prescription drug coverage - general information.

[39] C. for Medicare & Medicaid Services. Public use file: New data on geographic variation.

[40] C. for Medicare & Medicaid Services. Medicare enrollment reports, 2012.

[41] C. for Medicare & Medicaid Services. Medicare provider utilization and payment data: Physician and other supplier, 2012.

[42] K. F. Foundation. Medicare advantage: Ma-pd plan enrollment as a percent of total medicare population.

[43] K. F. Foundation. State health facts.

[44] K. F. Foundation. The facts on medicare spending and financing, July 2014.

[45] J. Fox and S. Weisberg. Robust regression in r. Appendix to book, September 2012.

[46] V. Gabrel, C. Murat, and A. Thiele. Recent advances in robust optimization. *European Journal of Operational Research*, 235(3):471–483, 2014.

[47] L. E. Ghaoui and H. Lebret. Robust solutions to least-squares problems with uncertain data. *SIAM J. Matrix Anal. Appl.*, 18(4):1035–1064, October 1997.

[48] P. B. Ginsburg. Fee-for-service will remain a feature of major payment reforms, requiring more changes in medicare physician payment. *Health Affairs*, 31(9):1977–83, 2012.

[49] J. Glazer and T. McGuire. Optimal risk adjustment in markets with adverse selection: An application to managed care. *The American Economic Review*, 90(4):1055–1071, 2000.

[50] J. Glazer, T. G. McGuire, and J. Shi. Risk adjustment of health plan payments to correct inefficient plan choice from adverse selection. Working Paper 19998, National Bureau of Economic Research, March 2014.

[51] D. Gregori, M. Petrinco, S. Bo, A. Desideri, F. Merletti, and E. Pagano. Regression models for analyzing costs and their determinants in health care: an introductory review. *International Journal for Quality in Health Care*, 23(3):331–341, April 2011.

[52] M. Hauskrecht and B. Kveton. Linear program approximations for factored continuous-state markov decision processes. In *Advances in Neural Information Processing Systems*, page None, 2003.

[53] G. Heo, B. Schmuland, and D. P. Wiens. Restricted minimax robust designs for misspecified regression models. *The Canadian Journal of Statistics*, 29(1):117–128, 2001.

[54] W. Hwang, J. Chang, M. LaClair, and H. Paz. Effects of integrated delivery system on cost and quality. *The American Journal of Managed Care*, 19(5):175–184, 2013.

[55] S. S. Iyengar and M. R. Lepper. When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology*, 79(6):995–1006, 2000.

[56] K. KB, L. MR, B.-P. C, Z. F, and W. JF. The impact of high-deductible health plans on men and women: an analysis of emergency department care. *Medical Care*, 51(8):639–45, 2013.

[57] J. P. Kuwornu, L. M. Lix, J. M. Quail, E. Wang, M. Osman, and G. F. Teare. A comparison of statistical models for analyzing episodes-of-care costs for chronic obstructive pulmonary disease exacerbations. *Health Serv Outcomes Res Method*, 13:203–218, 2013.

[58] R. G. Lane, S. M. Morrill, I. I. Rosen, and J. A.Wong. Dose volume considerations with linear programming optimization. *Med. Phys.*, 18(6):1201–1210, 1991.

[59] M. J. Laugesen and S. A. Glied. Higher fees paid to us physicians drive higher spending for physician services compared to other countries. *Health Affairs*, 30(9):1647–56, September 2011.

[60] J. E. Mason, B. T. Denton, S. A. Smith, and N. D. Shah. Using electronic health records to monitor and improve adherence to medication. Working paper, 2012.

[61] J. M. McWilliams, J. Hsu, and J. Newhouse. New risk-adjustment system was associated with reduced favorable selection in medicare advantage. *Health Affairs*, 31(12):2630–2640, 2012.

[62] B. Mihaylova, A. Briggs, A. O'Hagan, and S. G. Thompson. Review of statistical methods for analysing healthcare resources and costs. *Health Economics*, 20:897–916, 2011.

[63] D. C. Miller, C. Gust, J. B. Dimick, N. Birkmeyer, J. Skinner, and J. D. Birkmeyer. Large variations in Medicare payments for surgery highlight savings potential from bundled payment programs. *Health affairs (Project Hope)*, 30(11):2107–15, Nov. 2011.

[64] A. Nilim and L. E. Ghaoui. Robustness in markov decision problems with uncertain transition matrices. In S. Thrun, L. Saul, and B. Schölkopf, editors, *Advances in Neural Information Processing Systems 16*, pages 839–846. MIT Press, 2004.

[65] J. Rau. Methodology: How value-based purchasing payments are calculated. Technical report, Kaiser Health News, 2013.

[66] A. B. Rosen, A. Aizcorbe, A. J. Ryu, N. Nestoriak, D. M. Cutler, and M. E. Chernew. Policy makers will need a way to update bundled payments that reflects highly skewed spending growth of various care episodes. *Health affairs (Project Hope)*, 32(5):944–51, May 2013.

[67] S. B. Saxena and N. Holobinko. Why less choice is more in health insurance exchanges. *Harvard Business Review*, October 2013.

[68] A. J. Schaefer, M. D. Bailey, S. M. Shechter, and M. S. Roberts. *Operations Research and Health Care*, volume 70 of *International Series in Operations Research & Management Science*, chapter Modeling Medical Treatment Using Markov Decision Processes, pages 593–612. Springer US, 1 edition, 2004.

[69] S. A. Schroeder and W. Frist. Phasing out fee-for-service payment. *The New England Journal of Medicine*, 368:2029–32, May 2013.

[70] F. Schulte and D. Donald. How doctors and hospitals have collected billions in questionable medicare fees. The Center for Public Integrity, September 2012.

[71] P. Shi, J. J. Ye, and J. Zhou. Minimax robust designs for misspecified regression models. *The Canadian Journal of Statistics*, 31(4):397–414, December 2003.

[72] M. Smithson and J. Verkuilen. A better lemon squeezer? maximum-likelihood regression with beta-distributed dependent variables. *Psychological Methods*, 11(1):54–71, 2006.

[73] A. L. Soyster. Convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations Research*, 21:1154–1157, 1973.

[74] Staff. Value-based purchasing: a strategic overview for health care industry stakeholders. Technical report, Deloitte Center for Health Solutions, 2011.

[75] Staff. Department of Health and Human Services 45 CFR parts 153, 155, 156, 157 and 158. *Federal Register*, 77(236), 2012.

[76] Staff. Details for title: FY 2013 final rule tables. Technical report, Centers for Medicare and Medicaid Services, http://go.cms.gov/1sCr99T, 2013.

[77] Staff. ORYX risk adjustment guide. Technical report, Joint Commission, 2014.

[78] Staff. Profitability and revenue growth drop in US not-for-profit hospital preliminary medians. Technical report, Moody's Investors Service, 2014.

[79] C. J. Swearingen, M. S. M. Castro, and Z. Bursac. Modeling percentage outcomes: The %beta_regression macro. In *SAS Global Forum 2011*, number Paper 335-2011, 2011.

[80] A. A. Szpiro, K. M. Rice, and T. Lumley. Model-robust regression and a bayesian "sandwich" estimator. *The Annals of Applied Statistics*, 4(4):2099–2113, 2010.

[81] A. Tamar, S. Mannor, and H. Xu. Scaling up robust mdps using function approximation. In *Proceedings of The 31st International Conference on Machine Learning*, pages 181–189, 2014.

[82] R. C. Vliet. Predictability of individual health care expenditures. *The Journal of Risk and Insurance*, 59(3):443–461, September 1992.

[83] Washington State Hospital Association. *Manual: Integrated Delivery Systems*.

[84] J. Weiner, E. Trish, C. Abrams, and K. Lemke. Adjusting for risk selection in state health insurance exchanges will be critically important and feasible, but not easy. *Health Affairs*, 31(2):306–315, 2012.

[85] J. F. Wharam1, F. Zhang, B. E. Landon, S. B. Soumerai, and D. Ross-Degnan. Low-socioeconomic-status enrollees in high-deductible plans reduced high-severity emergency care. *Health Affairs*, 32(8):1398–1406, 2013.

[86] W. Wiesemann, D. Kuhn, and B. Rustem. Robust markov decision processes. *Mathematics of Operations Research*, 38:153–183, 2013.

[87] R. Winkelman and R. Damler. Risk adjustment in state Medicaid programs. *Health Watch*, 2008.

[88] R. Winkelman and S. Mehmud. A comparative analysis of claims-based tools for health risk assessment. Technical report, Society of Actuaries, April 2007.

[89] T. Xiao and A. C. Thiele. Analytical methods for plan efficiency evaluation in the health insurance exchanges. Technical report, Lehigh University, March 2015.

[90] H. Xu and S. Mannor. The robustness-performance tradeoff in markov decision processes. In B. Schölkopf, J. Platt, and T. Hofmann, editors, *Advances in Neural Information Processing Systems 19*, pages 1537–1544. MIT Press, 2006.

[91] R. Yi, J. Shreve, and W. Bluhm. Risk adjustment and its applications in global payments to providers. Milliman client report, Milliman,Inc., 2011.

[92] Y. Yih, editor. *Handbook of Healthcare Delivery Systems*. CRC Press, 1 edition, December 2010.

Biography

Tengjiao Xiao received her Bachelor of Economics in Finance from Sichuan University in June 2009. She also holds a M.S. degree in Analytical Finance from Lehigh University since May 2011. She is currently a PhD candidate in the Department of Industrial and Engineering at Lehigh University and is expected to graduate in May 2015. She has been a Rossin Doctoral Fellow since Fall 2014 and served the Lehigh University Student Chapter as treasurer during the academic year 2012-13. She spent the summer of 2013 working as Marketing Modeling Intern at Office Depot's headquarters in Boca Raton, FL. Since June 2014, she has been interning at Capital BlueCross in their Data Scientist team in Allentown, PA. She will be joining the Advisory Services practice at Ernst & Young in New York, NY as a Senior Consultant in June 2015.